



**UNIVERSIDAD
AUTÓNOMA DE MADRID**
Facultad de Ciencias
Depart. Química Orgánica



**CENTRO DE
INVESTIGACIONES
BIOLÓGICAS (CSIC)**

**Estudios sobre la conformación de quitoooligosacáridos
y glicomiméticos del antígeno Lewis A y su interacción
con dominios de heveína y galectinas. Una visión 3D
utilizando RMN.**

MEMORIA
que para optar al grado de
Doctor en Ciencias Químicas
Presenta

PALOMA VIDAL MORO
Madrid, 2011

Directores:
Jesús Jiménez Barbero
Juan Félix Espinosa Merino
Francisco Javier Cañada Vicinay

Agradecimientos

Detrás de cada logro siempre hay gente que de una manera u otra contribuyen a ello. Esta tesis no es una excepción y me gustaría agradecer a todas esas personas que lo han hecho posible.

En primer lugar, me gustaría agradecer a mis directores de tesis, Dr. Jesús Jiménez Barbero y Dr. Francisco Javier Cañada Vicinay, por permitirme el desarrollo de esta tesis en la lejanía y “a mi ritmo”.

Jesús, gracias por poner a mi disposición un proyecto que hiciera que esto de “a mi ritmo” fuera posible. Por tu paciencia, recibiendo resultados de tarde en tarde, por tus consejos y por tu apoyo durante todo este tiempo.

Quisiera hacer un agradecimiento muy especial a mi director de tesis, Dr. Juan Félix Espinosa Merino, por su ayuda, sus consejos, sus enseñanzas, su apoyo en los momentos bajos, su tesón... Juan Félix, gracias por el muchísimo tiempo que me has dedicado para discutir ideas y resultados. Por lo muchísimo que he aprendido contigo en estos años.

Quisiera también agradecer a los laboratorios Lilly, el haberme permitido utilizar los equipos de RMN para la realización de parte de este trabajo. En especial, quisiera agradecer al Dr. Alfonso Rivera Sagredo por todo su apoyo durante estos años.

Al Dr. Diego Cárdenas por haber aceptado ser el tutor de esta tesis y por su ayuda prestada.

Al grupo de la Dra. Sonsoles Martín Santamaría, por su colaboración en el capítulo 3 de esta tesis.

A Ángeles Canales, mi “estéreo”. Nuestros caminos se separaron pero siempre fuiste un gran apoyo, incluso en la lejanía... ¡gracias!

A mis compañeros de estos últimos años, que habéis compartido mi día a día. Especialmente a Aida, mi otra mitad en el laboratorio, y a Cris y a Pi por su apoyo y amistad.

Y finalmente a mi familia. A mis padres y mi hermano por todo su cariño y apoyo; por hacer de mí quien soy.

Y muy especialmente a Edu... sin ti esta tesis nunca hubiera visto la luz. Gracias por tu ayuda, activa y pasiva. Por aguantarme en mis momentos bajos, que no es fácil, y no dejar que nunca me rindiera. Esta es nuestra tesis... **¡al fin ha nacido!**

*A mis padres y a ti,
Edu, mi otra mitad.*

ÍNDICE

<u>Objetivos</u>	1
-------------------------------	----------

<u>Capítulo 1: Carbohidratos y Lectinas: Aplicación de la RMN a este tipo de moléculas</u>	5
1. Introducción y antecedentes	7
2. Aplicación de la RMN a Carbohidratos	8
1. Identificación de los diferentes sistemas de espín	
2. Asignación secuencial	
3. Determinación de la estructura tridimensional	
3. Aplicación de la RMN a Lectinas	16
1. Identificación de los diferentes sistemas de espín	
2. Asignación secuencial	
3. Asignación de NOEs a media y larga distancia	
4. Cálculo de estructuras	
4. Aplicaciones de la RMN al reconocimiento molecular de carbohidratos por proteínas	20
1. Métodos basados en la observación del ligando	
2. Métodos basados en la observación del receptor	
5. Bibliografía	26

<u>Capítulo 2: Estudio del β-C-Gal(1\rightarrow3)-β-Glc-OMe: comparativa entre el O- y C-disacárido</u>	29
1. Introducción y antecedentes	31
2. Análisis conformacional del disacárido β-Gal(1\rightarrow3)-β-Glc-OMe	38
2.1. Cálculos de mecánica molecular	
2.2. Estudio mediante RMN	
2.3. Conclusiones	
3. Análisis conformacional del disacárido β-C-Gal(1\rightarrow3)-β-Glc-OMe	45
3.1. Cálculos de mecánica molecular	
3.2. Estudio mediante RMN	
3.3. Conclusiones	
4. Conclusiones del análisis comparativo O/C-β-Gal(1\rightarrow3)β-Glc-OMe	54

5. Parte experimental	55
5.1. Materiales	
5.2. Mecánica Molecular	
5.3. Cálculos de constantes de acoplamiento y NOEs	
5.4. Experimentos de RMN	
6. Bibliografía	57

Capítulo 3: Estudio del β -C-Gal(1 \rightarrow 3)- β -Glc-OMe en presencia de galectina-1

	59
1. Introducción y antecedentes	61
2. Análisis conformacional del β-Gal(1\rightarrow3)β-Glc-OMe (1) en presencia de galectina-1	67
3. Análisis conformacional del β-C-Gal(1\rightarrow3) β-Glc-OMe (2) en presencia de galectina-1	69
4. Estudio de acoplamiento o <i>docking</i>	72
5. Conclusiones	78
6. Parte experimental	79
6.1. Materiales	
6.2. Estudio de Docking	
6.3. Experimentos de RMN	
7. Bibliografía	81

Capítulo 4: Estudio conformacional del Lewis a: comparativa entre el O- y C-trisacárido

	83
1. Introducción y antecedentes	85
2. Análisis conformacional Lewis a	89
2.1. Cálculos de mecánica molecular	
2.2. Estudio mediante RMN	
2.3. Conclusiones	
3. Análisis conformacional C-Lewis a	98
3.1. Cálculos de mecánica molecular	
3.2. Estudio mediante RMN	

3.3. Conclusiones	
4. Parte experimental	109
4.1. Materiales	
4.2. Mecánica Molecular	
4.3. Cálculos de constantes de acoplamiento y NOEs	
4.4. Experimentos de RMN	
5. Bibliografía	111

Capítulo 5: JAMFIS: Aplicación a los *C*-disacáridos. Comparación con los resultados de KISHI y con los *O*-disacáridos naturales113

1. Introducción y antecedentes	115
2. Método JAMFIS	117
3. Aplicación a <i>C</i>-disacáridos estudiados mediante mecánica molecular, NOEs y constantes de acoplamiento	121
3.1. <i>C</i> -lactosa	
3.2. α - <i>C</i> -manobiosa	
3.3. β - <i>C</i> -Gal-(1 \rightarrow 3)- β -Glc-OMe	
4. Aplicación a <i>C</i>-disacáridos analizados por Kishi	132
4.1. Enlace (1 \rightarrow 1): α,α -, β,β - y α,β - <i>C</i> -trehalosa	
4.2. <i>C</i> -disacáridos con uniones α (1 \rightarrow 3)	
4.3. <i>C</i> -disacáridos con uniones α (1 \rightarrow 4)	
4.4. <i>C</i> -disacáridos con uniones β -(1 \rightarrow 4)	
5. Conclusiones	156
6. Bibliografía	159

Capítulo 6: Dominios de Heveína: Mutantes del Ac-AMP2161

1. Introducción y antecedentes	163
2. Determinación de la estructura 3D del Ac-AMP2-F18W	168
2.1. Análisis de la proteína libre	
2.2. Análisis de la proteína asociada a (GlcNAc) ₃	
3. Determinación de la estructura 3D del Ac-AMP2-F18Nal	178
3.1. Análisis de la proteína libre	

3.2. Análisis de la proteína asociada a (GlcNAc) ₃	
4. Comparativa entre las mutantes y la proteína nativa Ac-AMP2	186
5. Estudios de interacción proteína-carbohidrato	187
6. Conclusiones	192
7. Parte experimental	193
7.1. Preparación de las muestras y experimentos de RMN	
7.2. Construcción de la librería del aminoácido no natural Nal	
7.3. Generación de estructuras 3D: DYANA y RMD (AMBER)	
7.4. Interacción proteína-carbohidratos: Valoraciones con Quitotriosa	
8. Bibliografía	198
<u>Conclusiones</u>	201

Objetivos



OBJETIVOS

Esta Tesis se puede dividir en dos partes bien diferenciadas:

1. Estudio conformacional de carbohidratos y análogos:

El objetivo general en este apartado ha sido continuar con nuestro estudio sistemático, utilizando técnicas de RMN y métodos de modelado molecular, de la comparación del comportamiento conformacional de *O* y *C*-disacáridos, naturales y miméticos, tanto en estado libre como asociado a proteínas. Como modelo, se ha escogido el disacárido Gal- β -(1 \rightarrow 3)-Glc- β -OMe, que es la unidad constituyente del antígeno sanguíneo Lewis a, y su análogo *C*-glicosídico. También se ha estudiado un oligosacárido más grande, en concreto el trisacárido Lewis a y su análogo *C*-glicosídico en el enlace β -(1 \rightarrow 3). Como modelo de proteína se ha elegido la galectina-1 humana.

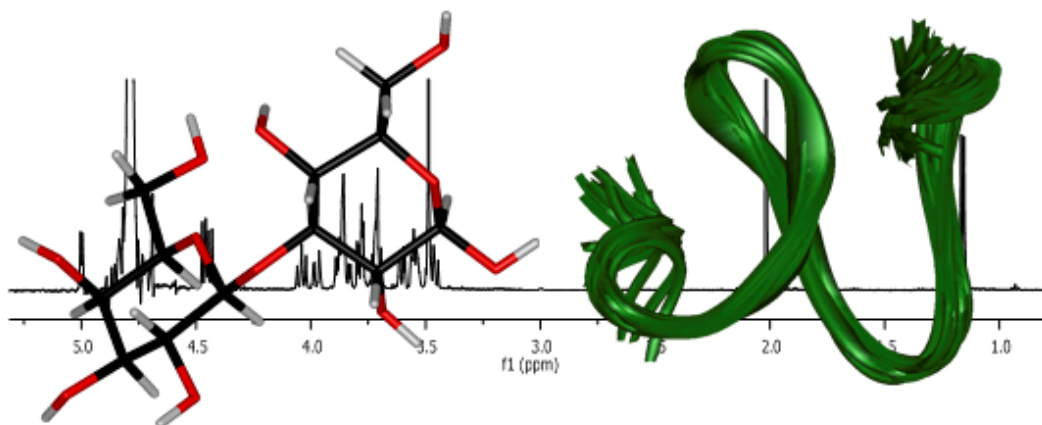
Desde el punto de vista metodológico, y también usando como modelos una variedad de análogos *C*-glicosídicos de los disacáridos naturales, nos propusimos comparar los resultados obtenidos mediante los métodos usuales de interpretación de datos de RMN con los deducidos mediante un análisis cuantitativo de datos experimentales de constantes de acoplamiento vecinales.

2. Interacciones proteína-carbohidrato.

En este apartado, y siguiendo con nuestra línea de trabajo en esta área, nos propusimos determinar la importancia de la naturaleza química de los anillos aromáticos presentes en los sitios de reconocimiento de lectinas para interaccionar con oligosacáridos. Como modelos, escogimos la lectina Ac-AMP2 y sus mutantes F18W y F18Nal, en los que se ha modificado el aminoácido en la posición 18 y, como oligosacárido, la *N*, *N'*, *N''*-triacetilquitotriosa (GlcNAc)₃.

Capítulo 1

Carbohidratos y Lectinas: Aplicación de la RMN a este tipo de moléculas.



1. Introducción y antecedentes:

La *Resonancia Magnética Nuclear* (RMN) es la técnica espectroscópica fundamental para la elucidación estructural de compuestos orgánicos e inorgánicos en disolución. Además, en los últimos 25 años, se ha demostrado inequívocamente que también permite la determinación de la estructura tridimensional de moléculas de características muy diferentes, incluyendo las moléculas de la vida: las biomoléculas¹⁻⁵. Pese a los avances, tanto conceptuales como de diseño de los espectrómetros, que han tenido lugar en los últimos años, es necesario enfatizar que el uso de la RMN puede estar restringido, dependiendo del tamaño de la molécula en cuestión. Aunque en los últimos años se ha incrementado mucho el rango de tamaños moleculares que es posible analizar por RMN, empleando marcajes isotópicos⁶⁻⁷ (¹³C, ¹⁵N, ²H) y metodologías de tipo TROSY⁸, todavía hoy no es posible, en general, utilizar indiscriminadamente la RMN para deducir la estructura 3D de proteínas o moléculas de gran tamaño. No obstante, sigue siendo posible conseguir información detallada de procesos de reconocimiento molecular mediante experimentos de RMN diseñados *ad hoc*. En general, los fenómenos de relajación son muy eficaces para moléculas de gran tamaño, con un número elevado de átomos de hidrógeno y esto produce un ensanchamiento dramático de las señales de RMN, con lo que resulta muy difícil, o imposible (incluso usando TROSY o CRINEPT⁹) conseguir los espectros de alta calidad necesarios para determinar la estructura 3D. En proteínas sin marcaje isotópico, como las que abordaremos aquí, el límite está en unos 10-15 kDa.

La difracción de rayos X es la técnica de referencia para realizar estudios estructurales a escala atómica. A pesar de que constituye una herramienta de incalculable valor en la determinación estructural tanto de moléculas pequeñas como de macromoléculas, su utilidad en algunos casos se ve limitada por varios motivos. Uno de los motivos que limita a esta técnica es la dificultad de obtener cristales en ciertas moléculas. Además, aunque proporciona información muy detallada, ésta se restringe a lo que sucede en una red cristalina dando una visión estática de la estructura de la molécula. En el caso del estudio de los carbohidratos, el objeto fundamental de esta tesis, al ser éstos moléculas flexibles, pueden adoptar varias conformaciones en disolución, por lo que la información aportada por esta técnica

puede ser simplemente parcial o no ser completa.

Dado que la mayoría de los procesos biológicos se desarrollan en disolución, el conocimiento de la estructura, dinámica e interacciones de las biomoléculas en disolución es de gran relevancia, y el uso de la RMN ha resultado ser ampliamente aceptado.

2. Carbohidratos:

Los carbohidratos constituyen uno de los más abundantes tipos de biomoléculas que se encuentran en la naturaleza. En las últimas décadas, se ha demostrado que los carbohidratos están involucrados en un gran número de funciones biológicas, desde la diferenciación celular hasta la regulación de patologías. Hoy en día, está claro que una de las principales funciones de los carbohidratos en los sistemas biológicos es la de ser puntos específicos de reconocimiento molecular¹⁰⁻¹¹. El reconocimiento específico de carbohidratos por otras biomoléculas (anticuerpos, enzimas, lectinas, ácidos nucleicos) es fundamental en las primeras etapas de muchos de esos procesos biológicos esenciales: respuestas inmunológicas, reconocimiento de moléculas propias o extrañas; invasión vírica y bacteriana; fertilización; adhesión celular; mecanismos de acción de antibióticos; metástasis tumorales etc. En este sentido, una de las claves para entender algunas de las mencionadas interacciones biológicas pasa por entender el papel que desempeñan los carbohidratos en dicha interacción.

Esta capacidad de los carbohidratos para actuar como marcadores en los procesos de reconocimiento molecular se debe a su enorme potencial para codificar la información biológica a través de su estructura espacial¹²⁻¹³. En lo referente a la conformación de los azúcares, los efectos estéricos y los efectos estereoelectrónicos son los factores fundamentales que determinan su estructura tridimensional. Entre estos últimos los más importantes son los denominados efecto “anomérico” y efecto “*exo*-anomérico”¹⁴⁻¹⁵.

Se denomina efecto anomérico a la tendencia del sustituyente electronegativo de la posición 1 del anillo de piranosa a colocarse en la posición axial, en contra de lo previsible a partir de las interacciones estéricas. Esta preferencia tiene su origen en la deslocalización hiperconjugativa del par electrónico no enlazante del oxígeno

endocíclico, O5, con el orbital σ^* del enlace contiguo C1-O1, que únicamente es posible en la orientación axial. Debido a esto, en la configuración axial (anómero α) se observa un acortamiento del enlace O5-C1, un alargamiento del enlace C1-O1 y un aumento del ángulo O5-C1-O1, con respecto al análogo ecuatorial. La magnitud del efecto anomérico varía con la naturaleza del grupo electronegativo, con la polaridad del disolvente y con la orientación y naturaleza química del resto de sustituyentes de la molécula (**Figura 1.1**).

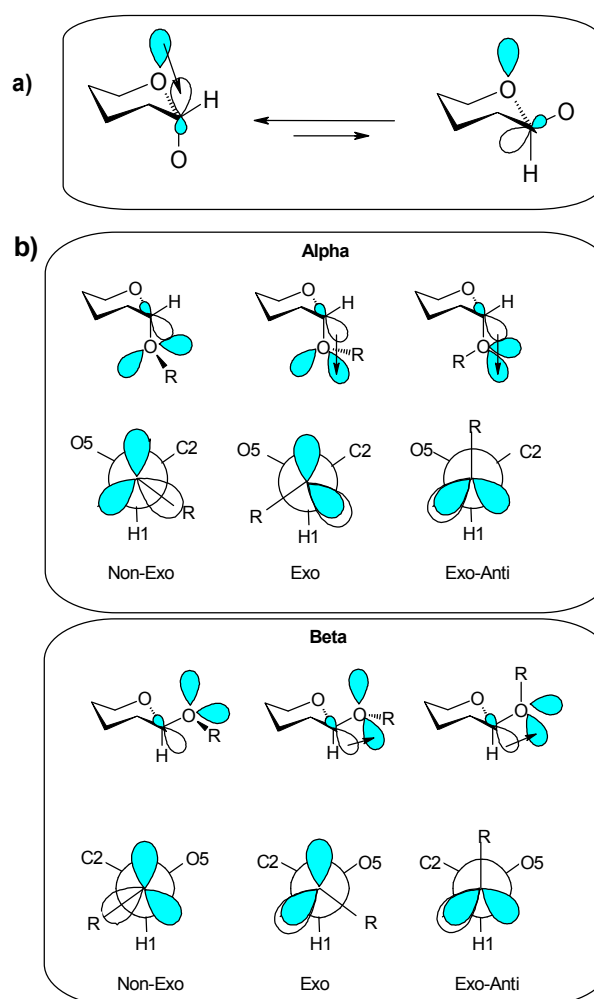


Figura 1.1: Representación esquemática del efecto anomérico (a) y *exo*-anomérico (b).

Por otro lado, el efecto *exo*-anomérico describe la preferencia que muestra la torsión exocíclica para adoptar una orientación *gauche*. De esta manera, la densidad electrónica del par solitario sobre O1 puede deslocalizarse respecto al σ^* del enlace C1-O5. Es decir, afecta al ángulo diedro definido por el enlace glicosídico y determina que aquellas conformaciones alrededor de este enlace que permitan la

interacción entre el par electrónico del O1 con el orbital σ^* del enlace endocíclico C1-O5 se encuentren favorecidas.

La combinación de interacciones estéricas y electrostáticas hacen que se favorezcan especialmente ciertas orientaciones del enlace glicosídico (ϕ : H1-C1-O1-Cx). En concreto, los ángulos -60° para las formas α y 60° para las formas β (**Figura 1.1**).

Desde el punto de vista de la RMN, hay que mencionar que, a pesar del tamaño relativamente pequeño de los oligosacáridos, especialmente cuando se comparan con otras biomoléculas, como proteínas y ácidos nucleicos, sus espectros de RMN son marcadamente complejos debido al gran solapamiento de las señales de la mayoría de los protones de los anillos. En general, solamente los protones anoméricos aparecen aislados a mayor desplazamiento químico. A partir de la medida de la constante de acoplamiento entre un protón anomérico y el protón adyacente, H2, es posible determinar la configuración anomérica de los distintos oligosacáridos. Lemieux *et al.*¹⁶ fueron los primeros en observar que la constante de acoplamiento vecinal entre protones en orientación axial-axial (7-9 Hz) es aproximadamente entre dos y tres veces mayor que en el caso de orientaciones axial-ecuatorial (<2 Hz), o ecuatorial-ecuatorial (4 Hz).

Estas reglas ayudan al proceso de identificación pero son necesarios experimentos adicionales para completar la asignación de las distintas resonancias del espectro. Este proceso consta en general de tres etapas:

1. *Identificación de los diferentes sistemas de espín:* Gracias a los acoplamientos escalares existentes entre los protones de una misma unidad de azúcar y a través de experimentos tipo COSY (**COR**relation **S**pectroscop**Y**)¹⁷ y TOCSY (**TOTAL** **COR**relation **S**pectroscop**Y**)¹⁸ es posible identificar las señales correspondientes a cada anillo. (**Figura 1.2**)

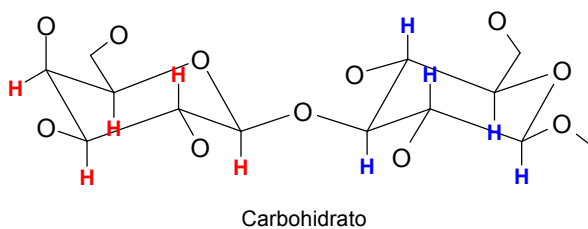


Figura 1.2: Esquema de un disacárido: el código de colores representa los dos sistemas de espín.

2. *Asignación secuencial:* La conexión secuencial entre los distintos anillos puede llevarse a cabo mediante varios métodos:

- *Experimentos heteronucleares:* El experimento **HMBC** (**H**eteronuclear **M**ultiple **B**ond **C**orrelation)¹⁹ permite determinar la conectividad ^1H - ^{13}C a dos y tres enlaces. Estudiando esta conexión es posible determinar la posición de unión de las unidades de azúcar. Para facilitar su análisis, previamente se puede adquirir el experimento **HSQC** (**H**eteronuclear **S**ingle **Q**uantum **C**orrelation)²⁰ que permite la asignación de los carbonos protonados mediante correlaciones entre el protón y el carbono al que se encuentra unido. La dispersión en la dimensión de carbono es mayor y facilita la asignación de los carbonos, y por extensión la de sus correspondientes protones.
- *Experimentos homonucleares:* A través del experimento **NOESY** (**N**uclear **O**verhauser **E**nhancement **S**pectroscopY)²¹ o **ROESY** (**R**otational nuclear **O**verhauser **E**ffect **S**pectroscopY)²² es posible establecer también la secuencia de los distintos residuos de una cadena oligosacáridica. En estos experimentos se observan picos de cruce entre los núcleos que se encuentran suficientemente cerca en el espacio.

3. *Determinación de la estructura tridimensional:* Una vez determinada la asignación secuencial, se procederá a la asignación de todos los picos de cruce del espectro **NOESY** o **ROESY**. En el caso de carbohidratos, el número de NOEs suele ser muy escaso, y por sí solos no son suficientes para determinar la estructura. Por este motivo es necesario llevar a cabo cálculos de mecánica molecular (**MM**) o de dinámica molecular (**DM**) que, en combinación con los datos experimentales de RMN, permiten definir completamente su estructura tridimensional²³. En esta tesis se optó por los cálculos de mecánica molecular.

- *Cálculos de MM:* La mecánica molecular se basa en la idea de que una molécula se puede representar por un conjunto de puntos con carga neta situados en los núcleos de los átomos y distribuidos en una superficie de energía potencial generada por ellos mismos. La energía total de la molécula se calcula como suma de energías de diferente naturaleza que están presentes

en la molécula:

$$U = \sum U_s + \sum U_b + \sum U_t + \sum U_{vdw} + \sum U_e$$

donde U es la energía potencial total del complejo o molécula que es la suma de la energía debida a la tensión de los enlaces individuales (*stretching*), U_s , de la energía de deformación de ángulo de enlace (*bending*), U_b , de la energía de torsión U_t , de la energía de las interacciones de Van der Waals U_{vdw} y de la energía de las interacciones electrostáticas U_e (**Figura 1.3**).

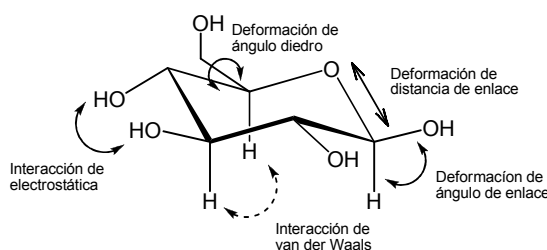


Figura 1.3: Representación de las interacciones presentes en una molécula.

La optimización de la geometría consiste en encontrar las coordenadas de una estructura molecular que represente un mínimo de energía potencial. Es, por tanto, muy importante que el campo de fuerzas elegido disponga de parámetros adecuados para la molécula en cuestión. El proceso de minimización de esta energía potencial en MM puede realizarse mediante diversos algoritmos (descenso pronunciado, gradientes conjugados,...). Estos algoritmos modifican la geometría de partida en sucesivas iteraciones hasta llevarla a un mínimo local cercano. No hay, en principio, forma de saber si se ha llegado al mínimo de energía absoluto o si, por el contrario, la estructura final es un mínimo local, sobre todo si se trata de moléculas con un alto número de grados de libertad. El resultado en estos cálculos es, en muchos casos, muy dependiente de la elección de la estructura inicial. Este problema se conoce habitualmente con el nombre de “problema de mínimos múltiples” y su solución pasa por utilizar algún método de búsqueda conformacional. (**Figura 1.4**)

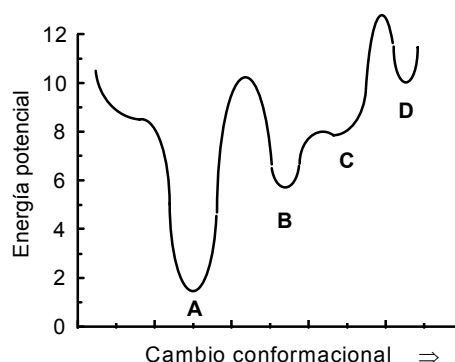


Figura 1.4. Representación esquemática del problema de mínimos múltiples.

Para realizar esta búsqueda conformacional, se pueden seguir dos tipos de aproximaciones: métodos de búsqueda estocástica, o métodos de búsqueda sistemática. La diferencia entre ambos métodos estriba en la forma de generar las estructuras de partida:

- a) Los métodos estocásticos generan muchas estructuras de partida utilizando variaciones aleatorias o pseudo-aleatorias de la geometría molecular. Generalmente, parten de geometrías estables y limitan la exploración del espacio conformacional a ciertas variaciones de coordenadas internas seleccionadas (ángulos de torsión, fundamentalmente). Los métodos aleatorios tienden a explorar todo el espacio conformacional si se deja que la búsqueda prosiga durante un número de iteraciones suficientemente largo.
- b) Los métodos sistemáticos, en cambio, exploran, de manera homogénea, todo el espacio conformacional. En el caso de los carbohidratos, suelen basarse en la suposición de que la estructura tridimensional de un oligosacárido viene principalmente determinada por las orientaciones de los enlaces glicosídico y aglicónico, por lo que la geometría de los conformeros de baja energía va a depender principalmente de las rotaciones en torno a dichos enlaces. Así, el espacio conformacional accesible a un determinado disacárido puede ser explorado eficazmente mediante la variación sistemática de dichas torsiones, rotando entre -180° y 180° con un incremento determinado. Una vez generado de esta forma el conjunto de conformeros que cubren todo el espacio conformacional,

el siguiente paso es obtener los mapas conformacionales. Se pueden definir tres tipos de mapas: mapas "rígidos", mapas "relajados" y mapas "adiabáticos". Los mapas rígidos de energía potencial se obtienen simplemente calculando la energía de cada uno de las conformaciones obtenidas mediante la expresión del campo de fuerzas seleccionado. Para generar los mapas relajados de energía potencial, se realiza una minimización de la energía de cada uno de los conformeros que forman el mapa rígido. Se permite la relajación energética de todos los grados de libertad, excepto las torsiones glicosídica y aglicónica. A partir de los mínimos de estos mapas se eligen conformeros representativos que se minimizan si ningún tipo de restricción para obtener las geometrías finales de los mismos. Por último, los mapas adiabáticos se generan a partir de diferentes mapas relajados en los que se han variado otros grados de libertad de la molécula (por ejemplo, el ángulo diedro que determina la orientación del hidroximetilo) eligiendo los puntos de menor energía de cada uno.

- *Combinación de los datos de RMN y de MM:* Una vez determinadas las regiones conformacionales de menor energía, estamos en disposición de analizar los datos de RMN obtenidos, especialmente los NOEs. Este análisis de NOEs se puede hacer de modo cualitativo o cuantitativo:
 - a) *Análisis cualitativo:* En los estudios de aplicación de RMN al análisis conformacional de carbohidratos, es habitual hacer un análisis de la disposición de los anillos alrededor de los enlaces glicosídicos haciendo uso de los denominados "NOEs exclusivos". De esta manera, se determina en qué regiones de este espacio conformacional una determinada pareja de protones se encuentra suficientemente próxima como para esperar un NOE entre ellos (hasta aproximadamente 4 Å en carbohidratos). Por tanto llamaremos "NOEs exclusivos" a aquellos NOEs que se observarían en una única región conformacional, ya que la distancia entre esos dos protones en el resto de regiones es demasiado grande para dar NOE. De esta forma, la observación de un NOE

exclusivo implica la existencia de la correspondiente familia conformacional, mientras que su ausencia permitiría descartar la participación significativa de dicha familia en el equilibrio conformacional.

- b) **Análisis cuantitativo:** Además de este estudio cualitativo, se puede realizar un estudio cuantitativo a través del análisis de la intensidad de los NOEs. Para ello, es importante tener en cuenta los factores de los que depende el NOE. El NOE es debido a la interacción dipolo-dipolo entre dos espines nucleares y, en el caso de dos protones aislados, la intensidad del NOE depende de la velocidad de relajación cruzada entre ambos núcleos que, a su vez, es función de las fluctuaciones en la orientación y la longitud del vector interprotónico. En el caso de moléculas pequeñas rígidas, en las que se asume que estos vectores se mueven isotrópicamente, las medidas de las intensidades de los NOEs entre pares de protones permiten obtener de forma directa distancias relativas, que pueden ser calibradas utilizando una distancia conocida (por ejemplo, el par intra-residuo H1-H3 en azúcares β). Esta aproximación es la denominada aproximación de espines aislados²⁴.

$$Int_{ref} * r_{ref}^6 = Int_{ij} * r_{ij}^6$$

Sin embargo, en presencia de otros protones, no existe relación directa sencilla entre la intensidad del pico y la constante de velocidad de relajación cruzada entre los dos protones, puesto que también puede contribuir la transferencia de magnetización a través de otros protones. Este hecho, junto con la posibilidad de que estén presentes movimientos internos, hace que las distancias obtenidas mediante la aproximación de espines aislados no sean exactas. Para tener en cuenta todos los protones se utiliza un método conocido como aproximación de la matriz de relajación completa²⁵⁻²⁶. En esta metodología, la matriz de las intensidades normalizadas de los picos de cruce (A) a un tiempo de mezcla dado (τ_m , el tiempo que se deja para que se produzca la

transferencia de magnetización entre protones) viene descrita por la siguiente expresión:

$$A = \exp[-R\tau_m]$$

$$R = \begin{bmatrix} \rho_i & \sigma_{ij} & \dots & \sigma_{in} \\ \sigma_{ji} & \rho_j & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \sigma_{ni} & \dots & \dots & \rho_n \end{bmatrix}$$

En esta expresión, R es la matriz de relajación completa. Cada elemento de la diagonal de esta matriz, ρ_i , es la constante de velocidad de autorelajación para un espín i, mientras cada elemento fuera de la diagonal, σ_{ij} , representa la constante de velocidad de relajación entre los espines i y j. Como estas constantes dependen de la inversa de la sexta potencia de la distancia (r_{ij}^{-6}) la matriz de relajación puede usarse para calcular NOEs teóricos a partir de distancias obtenidas mediante cálculos de mecánica o dinámica molecular.

3. Las lectinas:

Las lectinas son proteínas que unen carbohidratos con gran selectividad, reversibilidad y eficacia²⁷. Suelen aparecer en la superficie de las células regulando procesos de adhesión celular y, en algunos casos, ayudan al control del metabolismo celular. En general, el tamaño de estas proteínas impide una caracterización completa por RMN²⁸⁻²⁹. No obstante, existe una serie de lectinas de origen vegetal, de tamaño pequeño, que se enlazan a oligómeros de N-acetilglucosamina (GlcNAc), como el polisacárido quitina, componente principal del exoesqueleto de insectos y crustáceos y que está también presente en la pared celular de numerosos microorganismos. Estas lectinas se caracterizan por tener un motivo estructural común, denominado dominio de heveína. De hecho, sus características estructurales intrínsecas, así como su relativa fácil disponibilidad, han permitido su utilización como modelos de estudio para investigar el mecanismo, el origen y las características estructurales

involucradas en la formación de los complejos proteína-carbohidrato. De esta manera, se ha determinado la estructura molecular de distintos dominios de heveína, bien libres o bien unidos a carbohidratos, mediante RMN²⁹⁻³¹.

La estrategia general de RMN de proteínas consta de 4 etapas, de acuerdo con el protocolo propuesto por Wüthrich⁵:

1. *Identificación de los diferentes sistemas de espín*: La separación de los sistemas de espines pertenecientes a residuos de tipo AMX (*i.e.* Tyr, Trp, etc), de cadena larga (*i.e.* Lys, Arg, etc), o únicos (*i.e.* Ala, Val, etc), se consigue a partir de las conectividades entre protones de un mismo residuo debidas a acoplamientos escalares, observados en experimentos 2D COSY y/o TOCSY (**Figura 1.5**).

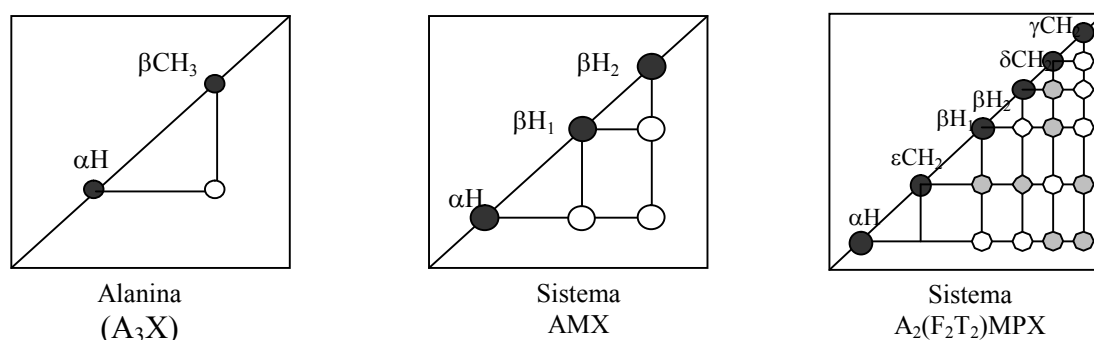


Figura 1.5: Diagrama de conectividades de algunos sistemas de espín típicos. Los puntos blancos representan picos de cruce COSY y TOCSY, mientras que los grises son exclusivamente picos de cruce TOCSY.

2. *Asignación secuencial*: Esta conexión se lleva a cabo a partir de los NOEs observados entre parejas de protones pertenecientes a residuos contiguos en la secuencia del péptido. Para ello se emplean experimentos 2D-NOESY a distintos tiempos de mezcla (**Figura 1.6**).

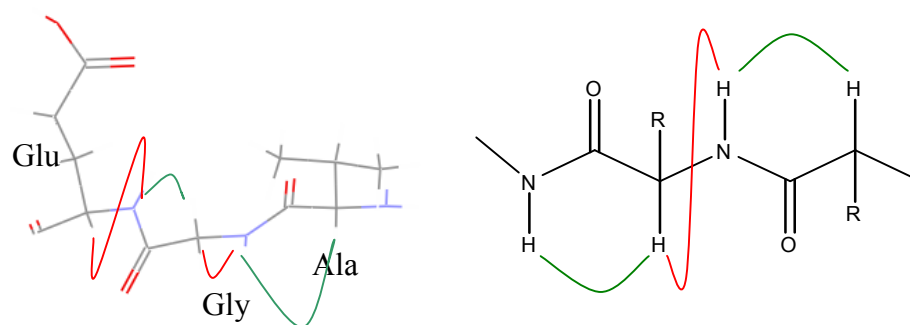


Figura 1.6: Ejemplo de NOEs secuenciales e intra-residuo. En verde se presentan NOEs entre residuos contiguos, mientras en rojo se muestran los NOEs intra-residuo.

3. *Asignación de NOEs a media y larga distancia:* Una vez asignados los NOEs intra-residuo y secuenciales, se asigna el resto de las señales del espectro NOESY. La intensidad, el número y el tipo de correlaciones NOESY a media distancia indicarán el tipo de estructura secundaria mientras que los NOEs a larga distancia permitirán determinar la estructura terciaria de la proteína (**Figura 1.7**).

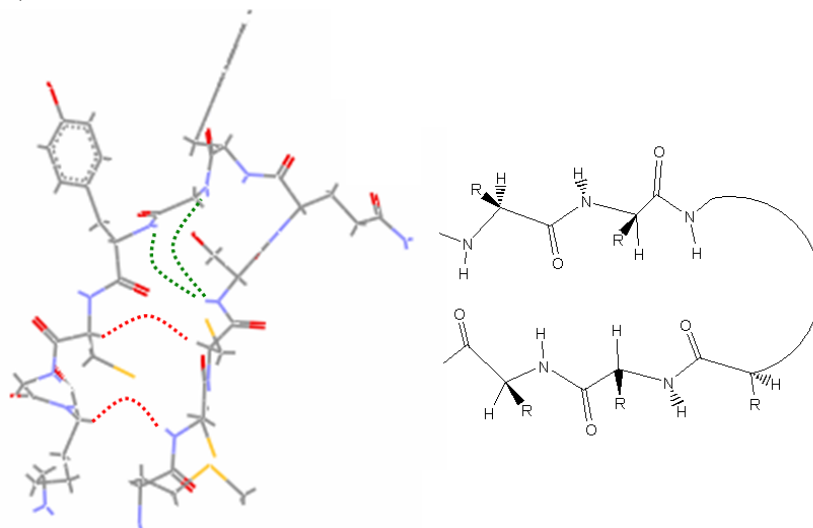


Figura 1.7: Ejemplo de NOEs a media y larga distancia. En verde se representan las correlaciones a media distancia, mientras en rojo se muestran aquellos residuos separados por más de 5 aminoácidos.

4. *Cálculo de estructuras:* Para determinar la estructura tridimensional de la proteína, es necesario convertir los NOEs en restricciones de distancias. Posteriormente, se emplean métodos computacionales que, además de incluir las restricciones de los NOEs, también consideran los parámetros intrínsecos de la estructura covalente de la proteína (longitudes de enlace, ángulos de enlace). Por ello, es necesario conocer la secuencia de aminoácidos. Cuantas más restricciones de distancia seamos capaces de identificar en el espectro NOESY, mejor podremos definir la estructura tridimensional en disolución. Existen diferentes métodos para determinar la estructura tridimensional en disolución a partir de restricciones de NOEs:
- *Métodos basados en geometría de distancias:* El primer paso para generar la estructura tridimensional es utilizar estas restricciones de distancia experimentales mediante un algoritmo matemático denominado “Geometría de Distancias”³²⁻³³. Este protocolo crea una matriz de distancias

interatómicas, que genera estructuras consistentes con las restricciones de distancia aplicadas. Estos métodos pueden funcionar en el espacio de coordenadas cartesianas o en el espacio de los ángulos de torsión, que es el más utilizado. En cualquier caso, el proceso parte de una estructura al azar, por ejemplo, con una cadena peptídica totalmente desplegada.

- Método DYANA (*DY*namics *A*lgorithm for *NMR* *A*pplications): El programa DYANA³⁴ se aplica para el cálculo de la estructura tridimensional de proteínas, a partir de las restricciones de distancia y ángulos de torsión obtenidos mediante RMN a partir de NOEs y constantes de acoplamiento respectivamente. DYANA se basa en una combinación de enfriamiento simulado y dinámica molecular en el espacio multidimensional de los ángulos de torsión, hasta encontrar la estructura 3D que cumpla las restricciones experimentales. Mediante la dinámica molecular restringida (*rMD*) los átomos de la estructura inicial están en movimiento hasta que el proceso alcanza un mínimo local. En principio, no existe ninguna garantía de que el mínimo alcanzado sea el mínimo global. Para intentar garantizar que se alcanza este mínimo global, primero se simula un calentamiento del sistema a alta temperatura, que posteriormente se irá enfriando gradualmente. El resultado de la primera simulación es una estructura de baja energía, normalmente un mínimo local. El proceso se repite de manera iterativa y cada estructura obtenida se usa nuevamente como inicial para el subsiguiente proceso de enfriamiento simulado. Finalmente, si la convergencia en el espacio conformacional es correcta, se alcanza el mínimo global (**Figura 1.8**).

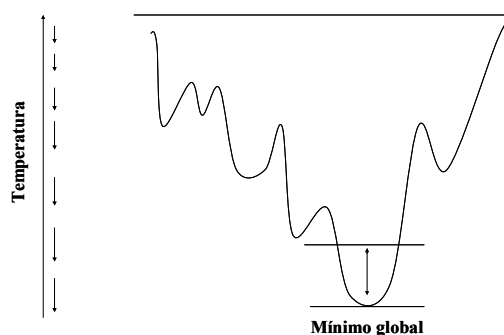


Figura 1.8: Representación del proceso para alcanzar el mínimo global en un proceso de enfriamiento simulado.

Finalmente, como el método DYANA no optimiza los contactos y enlaces del polipéptido, las mejores estructuras basándose en su función dyana se optimizarán mediante un protocolo de dinámica molecular restringida (RMD)³⁵, utilizando el campo de fuerzas AMBER³⁶. Este último proceso de refinamiento consiste en una minimización restringida con los NOEs experimentales, seguida de un protocolo de calentamiento y enfriamiento simulado, para finalmente terminar con otra minimización restringida.

4. Aplicaciones de la RMN al reconocimiento molecular de carbohidratos por proteínas:

El reconocimiento proteína-carbohidrato está involucrado en infecciones virales y en fenómenos de defensa de las plantas, entre otros muchos procesos.³⁷ Por este motivo, la elucidación de los mecanismos que gobiernan cómo los oligosacáridos se acomodan en los sitios de unión de lectinas, anticuerpos y enzimas es, hoy en día, motivo de amplia investigación.

Desde el punto de vista de la RMN, todos los parámetros espectroscópicos, como son los desplazamientos químicos, constantes escalares, difusión, velocidades de relajación, NOEs, etc, se afectan en mayor o menor medida, dependiendo de la afinidad y de la velocidad del intercambio entre los estados libre y asociado, al producirse la asociación ligando–receptor. Los métodos de RMN usados para la identificación y estudio de estas interacciones carbohidrato-proteína pueden dividirse en dos clases:

- Métodos que se basan en la observación de la variación de alguno de los parámetros de RMN del receptor motivadas por la interacción de éste con el ligando.
- Métodos que se basan en el seguimiento de los efectos producidos en los parámetros de RMN del ligando al pasar éste del estado libre al asociado.

1. *Métodos basados en la observación del receptor (proteína):* La observación de perturbaciones en señales del receptor previamente asignadas permite identificar el sitio de reconocimiento e identificar si se trata de una unión específica o no

específica. Este tipo de estudio permite, a su vez, caracterizar ligandos independientemente del régimen de intercambio químico. Las limitaciones de estos métodos basados en la observación del receptor residen en que se requiere que la proteína posea una serie de requisitos que no siempre se cumplen. La proteína debe ser accesible en cantidades suficientes, ser soluble y no estar agregada. En muchos casos, es necesario encontrar un sistema de expresión de la proteína que además permita llevar a cabo un marcaje isotópico eficaz (por ejemplo, ^{13}C , ^{15}N , ^2H), lo que posibilite la realización de diversos experimentos selectivos de RMN para facilitar la asignación de las señales. (**Figura 1.9**).

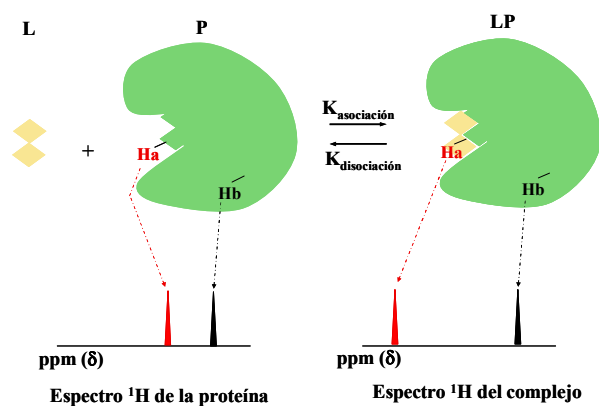


Figura 1.9: Representación de la asociación ligando-proteína mediante RMN. Se muestra la variación de desplazamiento químico para Ha (rojo), que se localiza en el sitio de reconocimiento, con respecto al cambio de desplazamiento químico de Hb (negro) que no se afecta.

Si el intercambio entre los estados libre y asociado es rápido en la escala de tiempos del desplazamiento químico, el seguimiento de los cambios en el desplazamiento químico puede permitir llevar a cabo, mediante valoraciones, la determinación de parámetros termodinámicos (constante de asociación (K_a), entalpía de unión, entropía, y la energía libre), usando una aproximación de *Van't Hoff*³⁸⁻⁴². En estos casos, las variaciones de desplazamiento químico de las señales de la proteína para las distintas concentraciones molares de ligando varían como se describe en la siguiente ecuación:

$$\delta_{obs} - \delta_{libre} = (\delta_{comp} - \delta_{libre}) \frac{[PL]}{P_T};$$

donde $[PL]$ es la concentración del complejo proteína–ligando, P_T es la concentración total de proteína, δ_{libre} y δ_{comp} son los desplazamientos químicos de la proteína libre y del complejo, respectivamente; δ_{obs} es el desplazamiento químico observado para una concentración dada de ligando. Puesto que la K_a es directamente proporcional a la concentración de complejo e inversamente proporcional al producto de las concentraciones de la proteína y ligando, la ecuación queda definida:

$$\delta_{\text{obs}} - \delta_{\text{libre}} = (\delta_{\text{comp}} - \delta_{\text{libre}}) \frac{K_a [P_T] [L_T]}{P_T}$$

Donde la concentración de complejo en disolución obedece a la siguiente ecuación:

$$[PL] = \frac{1}{2} \left\{ \left([P_T] + [L_T] + \frac{1}{K_a} \right) - \sqrt{\left([P_T] + [L_T] + \frac{1}{K_a} \right)^2 - 4[P_T][L_T]} \right\}$$

Por tanto, las variaciones en los desplazamientos químicos observados pueden ser utilizadas, de modo sencillo, para el cálculo de las constantes de asociación. Estos experimentos se realizan manteniendo constante la concentración de la proteína P_T , adquiriendo espectros de ^1H para cada una de las diferentes concentraciones de ligando. La determinación de la K_a se basa en la representación de la curva de valoración δ_{obs} frente a $[L]$, lo que permite realizar un ajuste no lineal a la ecuación, considerando que los parámetros variables son K_a y $(\delta_{\text{libre}} - \delta_{\text{comp}})$.

2. *Método basado en la observación del ligando (Carbohidrato):* Los métodos basados en la observación de las señales de los ligandos se apoyan en el hecho de que es posible distinguir entre los estados libre y asociado desde el punto de vista de la RMN, gracias a las diferencias de la velocidad rotacional y traslacional de una molécula pequeña al pasar del estado libre al asociado. La ventaja de este método radica en evitar la necesidad de producir grandes cantidades del receptor marcado isotópicamente. Además, tampoco es necesaria la asignación de la proteína, ni imprescindible el conocimiento de su estructura tridimensional. La ausencia de estos requisitos hace que estos métodos sean más

rápidos y de mayor aplicabilidad al estudio de los procesos de formación de los complejos que los basados en la observación del receptor. Obviamente, la información relativa al receptor que se puede obtener con estos métodos es más limitada que en los anteriores, aunque, en casos favorables, pueden complementarse eficazmente con protocolos de *docking*.

Las aproximaciones basadas en la observación del ligando se basan en la transferencia, mediada por intercambio químico, de la información del estado asociado al estado libre. Este requisito hace que estos métodos sean útiles para la identificación de ligandos con afinidades entre débiles y moderadas (intercambio relativamente rápido, con K_{off} mayor de $2\text{-}3\text{ s}^{-1}$).

Generalmente, los ligandos son moléculas de bajo peso molecular que se caracterizan por tener tiempos de correlación cortos lo que se traduce en velocidades de relajación lentas y coeficientes de difusión translacional relativamente altos. Teniendo en cuenta la relación que existe entre el efecto nuclear Overhauser y el tiempo de correlación (**Figura 1.10**) los picos de cruce de los experimentos NOESY de moléculas pequeñas presentan signos opuestos a los de la diagonal (NOEs positivos). Por el contrario, las moléculas de alto peso molecular como las proteínas, tienen tiempos de correlación largos, que se corresponden con velocidades de relajación rápidas y velocidades de difusión lentas. En este caso, los picos de cruce de los espectros NOESY de las moléculas de alto peso molecular aparecen con el mismo signo que los de la diagonal (NOEs negativos).

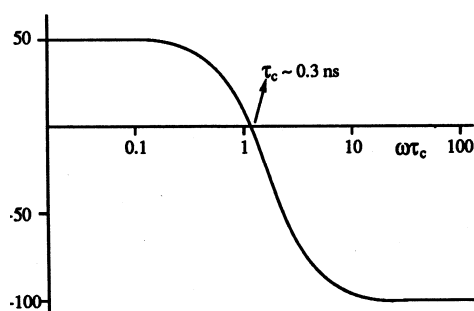


Figura 1.10: Dependencia del incremento del NOE homonuclear con $\omega\tau_c$. ω es la frecuencia del espectrómetro y τ_c el tiempo de correlación de la molécula.

Cuando una molécula pequeña (ligando) se asocia a una macromolécula, sus propiedades de RMN cambian y adquieren las propiedades de ésta. Por lo tanto,

en el estado asociado, los ligandos presentan velocidades de relajación rápidas, picos de cruce NOESY de signo igual que los de la diagonal y coeficientes de difusión más pequeños. Estas diferencias de comportamiento permiten realizar diversos experimentos de RMN capaces de diferenciar los estados libre y asociado del ligando y por tanto determinar si existe interacción entre el ligando y la macromolécula.

El cambio de signo en el NOE de un ligando cuando es reconocido por una proteína debido al cambio en el tiempo de correlación se conoce como NOE transferido (tr-NOE) y fue inicialmente descrito por Bothner-By⁴³ y más tarde utilizado por Feeney y sus colaboradores⁴⁴, así como por Clore y Gronenborn⁴⁵⁻⁴⁶.

Si se consideran un par de núcleos, I y S, de un ligando C que se encuentra en intercambio rápido en la escala de desplazamientos químicos entre la forma libre (L) y la unida (U) a un receptor R, el NOE observado en una señal determinada es un promedio resultante de la contribución de los NOEs de ambos estados, que se expresa en función de la constante de velocidad de relajación promedio ($\langle\sigma_{IS}\rangle$) y que viene descrita por la siguiente expresión:

$$\langle\sigma_{IS}\rangle = N^L\sigma_{IS}^L + N^U\sigma_{IS}^U$$

Siendo σ_{IS}^L y σ_{IS}^U las constantes de relajación cruzada de la forma libre y unida para el sistema IS, y N^L y N^U las fracciones molares correspondientes. Así, si en esta expresión domina el término de la forma unida, los NOEs observados caracterizarán a la conformación del ligando en el estado asociado, mientras que si el término que domina la expresión es el correspondiente al estado libre, los NOEs reflejarán la conformación del ligando en dicho estado. Por lo tanto, el experimento tr-NOESY es aplicable en los casos donde se cumpla:

$$|N^L\sigma_{IS}^L| \ll |N^U\sigma_{IS}^U|$$

La constante de velocidad de relajación cruzada (σ) está relacionada con el tiempo de correlación (τ_c). Las moléculas de peso molecular bajo tienen un τ_c corto y, como consecuencia, exhiben NOEs positivos. Cuando una molécula

pequeña se une a una proteína, su tiempo de correlación efectivo varía, y gira con movimientos lentos, como si fuera parte de la proteína, lo que da lugar a los NOEs transferidos negativos. Las condiciones para que el tr-NOE se produzca serán más favorables cuanto más grande sea la proteína, ya que el tamaño del complejo será mayor y, por tanto, el tiempo de correlación correspondiente al estado asociado del ligando también lo será.

Respecto a la afinidad del ligando, como se ha mencionado anteriormente, es necesario que exista un intercambio rápido entre la forma libre y la forma unida, de tal modo que se produzca transferencia de información entre los dos estados. Para que esto ocurra, la velocidad de disociación del complejo debe ser rápida en la escala de tiempos de relajación ($>2\text{-}3\text{ s}^{-1}$), lo que suele ocurrir en el rango de afinidades que va desde 100 nM a 10 mM.

Experimentalmente el tr-NOESY se realiza mediante una secuencia NOESY normal aplicada a un ligando en intercambio entre su estado libre y asociado. El ligando está en exceso respecto al receptor, por lo que los desplazamientos químicos observados están muy próximos a los del ligando libre. En general, si el intercambio es rápido en la escala de tiempos de la relajación, la existencia de interacción ligando-receptor puede detectarse fácilmente mediante inspección visual de los picos de cruce, y a partir de estos se puede derivar la conformación del ligando en el estado asociado (conformación bioactiva).

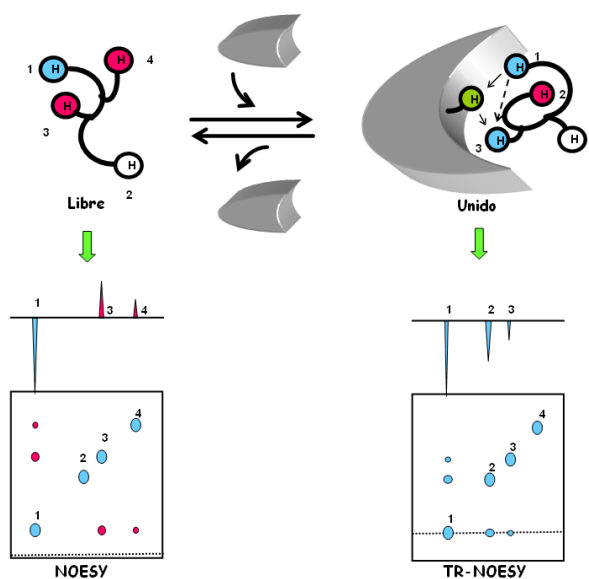


Figura 1.11: Esquema del experimento tr-NOESY. Los picos de cruce del ligando en el espectro son del mismo signo que los de la diagonal.

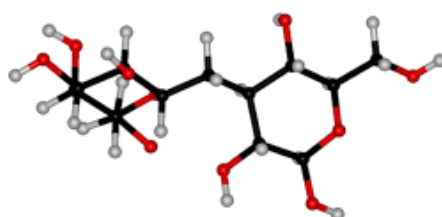
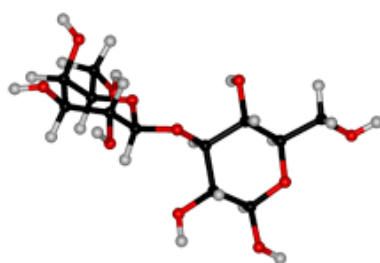
5. Bibliografía:

- (1) Bax, A.; Grezeiek, S. *Accounts Chem. Res.* **1993**, *26*, 131.
- (2) Guntert, P. Q. *Rev. Biophys* **1998**, *31*, 145.
- (3) Herrmann, T.; Guntert, P.; Wuthrich, K. *J. Biomol. NMR* **2002**, *24*, 171.
- (4) Wemmer, D. E.; Reid, B. R. *Annu. Rev. Phys. Chem.* **1985**, *36*, 105.
- (5) Wuthrich, K. *NMR of Proteins and Nucleic Acids* John Wiley&Sons. NewYork, 1986.
- (6) Kay, L. E.; Gardner, K. H. *Curr. Opin. Struct. Biol.* **1997**, *7*, 564.
- (7) Wagner, G. *J. Biomol. NMR* **1993**, *3*, 375.
- (8) Pervushin, K.; Riek, R.; Wider, G.; Wüthrich, K. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 12366.
- (9) Riek, R.; Wider, G.; Pervushin, K.; Wüthrich, K. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 4918.
- (10) Harlan, J. M.; Liu, D. *Adhesion: It's Role in Inflammatory disease*; W.H. Freeman N.Y, 1992.
- (11) Sharon, N.; Lis, H. *Curr. Opinion Struct. Biol.* **1991**, *1*, 741.
- (12) Gabius, H. J.; Siebert, H. C.; Andre, S.; Jimenez-Barbero, J.; Rudiger, H. *Chembiochem* **2004**, *5*, 740.
- (13) Laine, R. A. *Glycobiology* **1994**, *4*, 759.
- (14) Schreiner, P. R. *Angew Chem Int Edit* **2002**, *41*, 3579.
- (15) Tvaroska, I.; Bleha, T. *Adv Carbohydr Chem Bi* **1989**, *47*, 45.
- (16) Lemieux, R. U.; Kullnig, R. K.; Bernstein, H. J.; Schneider, W. G. *J Am Chem Soc* **1958**, *80*, 6098.
- (17) Aue, W. P.; Bartholdi, E.; Ernst, R. R. *J Chem Phys* **1976**, *64*, 2229.
- (18) Braunschweiler, L.; Ernst, R. R. *J Magn Reson* **1983**, *53*, 521.
- (19) Hurd, R. E.; John, B. K. *J. Magn. Reson.* **1991**, *91*, 648.
- (20) Willker, W.; Leibfritz, D.; Kerssebaum, R.; Bermel, W. *Magn Reson Chem* **1993**, *31*, 287.
- (21) Neuhaus, D. W., M. P. *The Nuclear Overhauser Effect in Structural and Conformational Analysis*; VCH: New York, 1989.
- (22) Bothnerby, A. A.; Stephens, R. L.; Lee, J. M.; Warren, C. D.; Jeanloz, R. W. *J Am Chem Soc* **1984**, *106*, 811.
- (23) Peters, T.; Pinto, B. M. *Curr Opin Struc Biol* **1996**, *6*, 710.
- (24) Gronenborn, A. M.; Clore, G. M. *Prog Nucl Mag Res Sp* **1985**, *17*, 1.
- (25) Boelens, R.; Koning, T. M. G.; Kaptein, R. *J Mol Struct* **1988**, *173*, 299.
- (26) Boelens, R.; Koning, T. M. G.; Vandermarel, G. A.; Vanboom, J. H.; Kaptein, R. *J Magn Reson* **1989**, *82*, 290.
- (27) Rudiger, H.; Siebert, H. C.; Solis, D.; Jimenez-Barbero, J.; Romero, A.; von der Lieth, C. W.; Diaz-Marino, T.; Gabius, H. J. *Curr. Med. Chem.* **2000**, *7*, 389.
- (28) *NMR Spectroscopy of Glycoconjugates*; Jimenez-Barbero, J., Peters, T., Ed.; Wiley-VCH, 2002.
- (29) Poveda, A.; Jimenez-Barbero, J. *Chem Soc Rev* **1998**, *27*, 133.
- (30) Asensio, J. L.; Canada, F. J.; Bruix, M.; Rodriguez-Romero, A.; Jimenez-Barbero, J. *Eur. J. Biochem.* **1995**, *230*, 621.
- (31) Jiménez-Barbero, J.; Javier Cañada, F.; Asensio, J. L.; Aboitiz, N.; Vidal, P.; Canales, A.; Groves, P.; Gabius, H.-J.; Siebert, H.-C. In *Advances in*

- Carbohydrate Chemistry and Biochemistry*; Derek, H., Ed.; Academic Press: 2006; Vol. Volume 60, p 303.
- (32) Brunger, A. T.; Nigles, M. *Quart. Rev. Biophys.* **1993**, 26, 49.
 - (33) Havel, T. F.; Wüthrich, K. *Bull. Math. Biol.* 1984, 46, 673-698 **1984**, 46, 673.
 - (34) Guntert, P.; Mumenthaler, C.; Wuthrich, K. *J. Mol. Biol.* **1997**, 273, 283.
 - (35) Scheek, R. M.; Van Gunsteren, W. F.; Kaptein, R. *Methods Enzymol* 1989, 177, 204.
 - (36) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. J. *J. Am. Chem. Soc.* **1984**, 106, 765.
 - (37) Dwek, R. A. *Chem. Rev.* **1996**, 96, 683.
 - (38) Asensio, J. L.; Canada, F. J.; Bruix, M.; Rodriguez-Romero, A.; Jimenez-Barbero, J. *Eur J Biochem* **1995**, 230, 621.
 - (39) Feeney, J.; Batchelor, J. G.; Albrand, J. P.; Roberts, G. C. K. *J. Magn. Reson.* **1979**, 33, 519.
 - (40) Kronis, K. A.; Carver, J. P. *Biochemistry* **1985**, 24, 834.
 - (41) Lian, L. Y.; Roberts, G. C. K. *NMR of Macromolecules* Oxford University Press, Oxford, 1993.
 - (42) Poveda, A.; Jiménez-Barbero, J. *Chem. Soc. Rev.* **1998**, 27, 133.
 - (43) Bothnerb.Aa; Gassend, R. *Ann Ny Acad Sci* **1973**, 222, 668.
 - (44) Albrand, J. P.; Birdsall, B.; Feeney, J.; Roberts, G. C. K.; Burgen, A. S. V. *Int J Biol Macromol* **1979**, 1, 37.
 - (45) Clore, G. M.; Gronenborn, A. M. *J Magn Reson* **1982**, 48, 402.
 - (46) Clore, G. M.; Gronenborn, A. M. *J Magn Reson* **1983**, 53, 423.

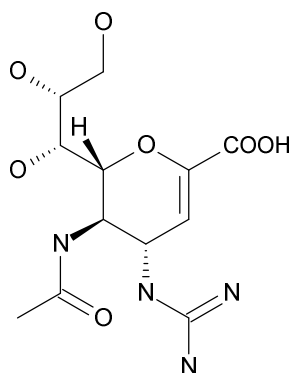
Capítulo 2

**Estudio del β -C-Gal(1 \rightarrow 3)- β -Glc-OMe:
comparativa entre el *O*- y el *C*-disacárido.**

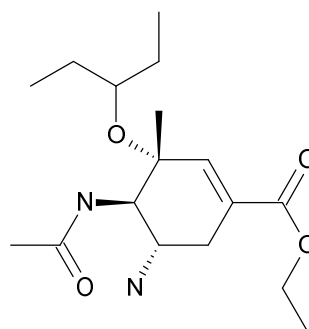


1. Introducción y antecedentes:

Los enlaces glicosídicos son relativamente inestables en distintas condiciones experimentales, especialmente en medios ácidos. Además, son susceptibles al ataque de enzimas glicosidasas, por lo que la alternativa de utilizar glicomiméticos como inhibidores de glicosidasas resulta especialmente atractiva. Su uso podría ayudar al desarrollo de fármacos con mayor estabilidad que el compuesto natural, pero manteniendo e incluso mejorando las propiedades de reconocimiento molecular¹⁻⁶. Este interés en glicomiméticos ha llevado al desarrollo de metodología sintética para poder disponer de una variedad de análogos glicosídicos en cantidades suficientes para poder realizar estudios biológicos. En la actualidad, ya existen algunos ejemplos de fármacos derivados de los carbohidratos como, por ejemplo, los antigripales Relenza® o Tamiflu®, o las heparinas modificadas.



Relenza® (Zanamivir)



Tamiflu® (Oseltamivir)

Figura 2.1: Estructuras de los fármacos Relenza® y Tamiflu® basados en estructuras de carbohidratos.

Los C-glicósidos son un tipo de glicomiméticos donde el átomo de oxígeno interglicosídico ha sido sustituido por un grupo metileno. Realizando una búsqueda en la base de datos de MDDR (*MDL drug data report*), que incluye los fármacos comercializados y los que están en desarrollo, observamos que existen varios ejemplos de C-glicósidos que se encuentran en fase III y muchos otros ejemplos que se encuentran en estadios más tempranos. Principalmente, son compuestos antibióticos, antimitóticos, antineoplásicos, inhibidores de la absorción del colesterol, compuestos terapéuticos contra la aterosclerosis o compuestos antilipídicos. En la siguiente tabla (**Tabla 2.1**) se han recogido algunos ejemplos.

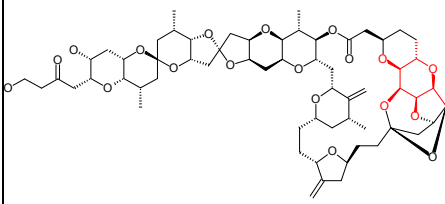
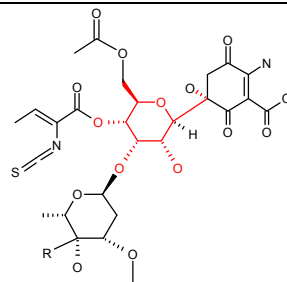
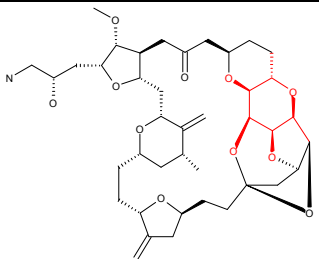
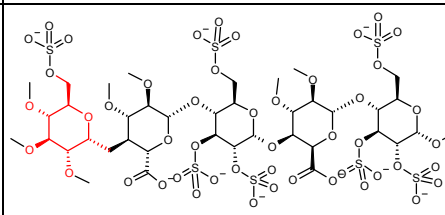
Nombre	Estructura	Tipo	Fase clínica
Isohomohalichondrin B (Pharmamar)		Antineoplásico Antibiótico Antitumoral	Preclínica
Paulomycin (Pfizer)		Antibiótico	Test biológicos
Eribulin (Eisai)		Antitumoral	Fase III
(Sanofi-Aventis)		Anticoagulante, Inhibidor del Factor Xa	Test biológicos

Tabla 2.1: Ejemplos de algunos derivados de C-glicósidos en distintas fases clínicas.

La sustitución del oxígeno provoca un cambio drástico en las características estructurales de la molécula, tanto en el tamaño como en las propiedades electrónicas de la unión glicosídica. Desde el punto de vista conformacional, esto afecta al efecto *exo-anomérico*. En los C-disacáridos, este efecto desaparece debido a la ausencia del oxígeno glicosídico. Este hecho es muy significativo, ya que el comportamiento conformacional de los glicósidos está gobernado en parte por el efecto *exo-anomérico* y puede implicar diferencias conformacionales importantes respecto al de los carbohidratos.

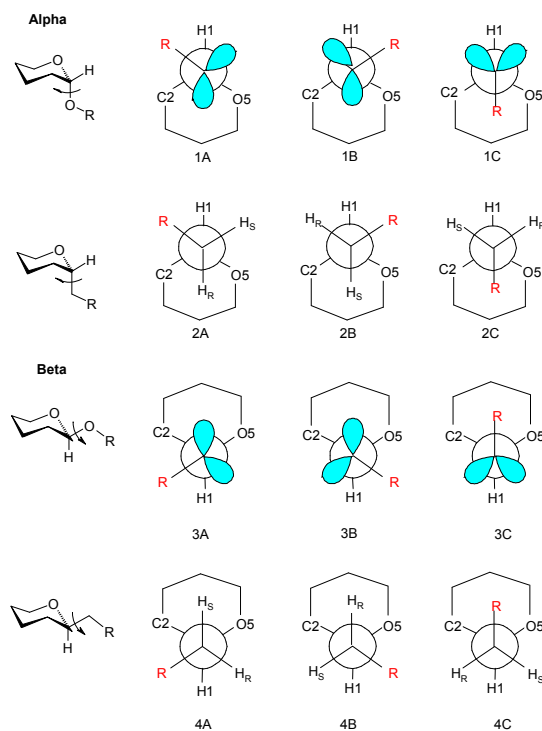


Figura 2.2: Conformaciones alternadas alrededor del enlace glicosídico para los *O*- y *C*- glicosídeos.

La primera comparación entre las estructuras tridimensionales de *C*- y *O*-glicosídeos fue realizada en 1987 por el grupo de Kishi⁷. En este primer trabajo, se estudió la orientación del enlace aglicónico en α y β -*C*-glicosídeos y se comparó con los correspondientes *O*-glicosídeos (**Figura 2.2**). En los *O*-glicosídeos, de las tres conformaciones alternadas posibles, dos se encuentran estabilizadas por el efecto *exo*-anomérico, 1(3)B y 1(3)C, ya que el orbital sp^3 se encuentra *anti*-periplanar al enlace C1-O5. De estas dos geometrías estabilizadas estéreo-electrónicamente, la conformación 1(3)B es la preferida, ya que la conformación 1(3)C sitúa al sustituyente R con orientación *gauche* a C2 y O5 y está más impedida estéricamente. En los *C*-glicosídeos, gracias a la presencia de los dos protones en el carbono pseudo-glicosídico, es posible medir las constantes de acoplamiento 1H - 1H a tres enlaces y, de esta manera, obtener información acerca de la conformación en torno a este enlace. Los valores de 3J obtenidos para distintos grupos R en α -derivados oscilaron entre 10 y 12 Hz para J_{H1-HpS} y entre 3 y 4 Hz para J_{H1-HpR} . Estos valores indican que, en la conformación mayoritaria, H1 se encuentra *anti*-periplanar a HpS y *gauche* respecto a HpR y que, por tanto, la conformación mayoritaria adoptada por los α -*C*-glicosídeos (confórmero 2B) es la misma que en su *O*-análogo, a pesar de no existir

ninguna estabilización estéreo-electrónica. El análisis cualitativo de los datos experimentales, en los β derivados dio lugar a conclusiones similares.

Estos resultados llevaron a Kishi y a sus colaboradores a sugerir que el efecto *exo*-anomérico no era el factor más importante que determinaba las características conformacionales de los oligosacáridos. Para comprobar esta hipótesis, su grupo de investigación llevó a cabo el análisis por RMN, principalmente mediante constantes de acoplamiento protón-protón, de un gran número de *C*-di y *C*-trisacáridos⁸⁻²⁰ para concluir que estos compuestos presentaban las mismas características conformacionales que los *O*-glicósidos en disolución. La metodología utilizada por Kishi consistía en situar a la molécula en una red de diamante y, teniendo en cuenta las conformaciones alternadas en torno a los enlaces interglicosídicos (**Figura 2.3**), deducir la conformación preferida a partir de las constantes de acoplamiento. Estos estudios pusieron de manifiesto que el enlace pseudo-glicosídico adoptaba la conformación *exo*-anomérica igual que en los *O*-glicósidos, mientras que el enlace pseudo-aglicónico adoptaba una conformación alternada en ausencia de interacciones estéricas tipo 1,3-diaxiales.

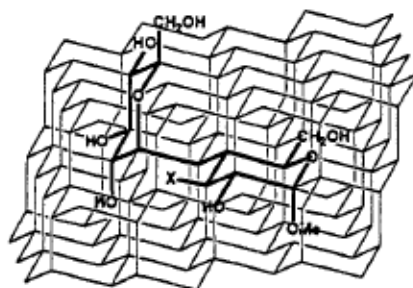


Figura 2.3: Representación de un *C*-disacárido superpuesto en una red de diamante.

Sin embargo, estudios de Rayos-X de tres *C*-disacáridos, en concreto de la *C*-sacarosa¹⁴, la *C*-isomaltosa²¹ y la *C*-gentibiosa²², pusieron de manifiesto que sus estructuras en estado sólido eran diferentes a la de los compuestos naturales. Estas diferencias estructurales, junto con las observadas previamente en disolución para otro tipo de glicomiméticos, los *S*-glicósidos, que son mucho más flexibles que los correspondientes *O*-análogos²³⁻²⁴, pusieron en duda que las conclusiones de Kishi de la similitud entre las propiedades conformacionales de los *C*- y *O*-glicósidos tuvieran

validez general y llevaron a la realización de estudios conformacionales de otros *C*-disacáridos mediante la combinación de RMN y cálculos de mecánica molecular.

El primero de estos estudios fue el análisis conformacional comparativo de la *C*-y *O*-lactosa, realizado en nuestro grupo de investigación²⁵⁻²⁶. Un estudio riguroso y sistemático sobre el disacárido α -metil-lactósido demostró que este es prácticamente monoconformacional en disolución acuosa, con aproximadamente un 90% de la población concentrada en la región denominada *syn* (zona central del mapa conformacional con el enlace glicosídico comprendido entre 0-90° y el aglicónico entre -60° y +60°), y sólo un 10% de la población localizada en la región denominada *anti- ψ* (región con el enlace glicosídico comprendido entre 0-90° y centrada en 180° en torno al enlace aglicónico). Por el contrario, el análisis conformacional de la *C*-lactosa, mediante cálculos de mecánica molecular indicó que, si bien la disposición adoptada en torno al enlace glicosídico era preferentemente *exo*-anomérica, como ocurría en la *O*-lactosa, en cambio, la disposición en torno al enlace aglicónico era marcadamente diferente. De hecho, mientras que en el análogo natural el mínimo global reside en la región *syn*, en la *C*-lactosa el mínimo global se encuentra en la región *anti- ψ* , concentrando un 55% de la población. El 45% restante se localiza predominantemente en la región *syn* aunque existe una pequeña población, en torno al 5%, en la región denominada *anti- ϕ* (región centrada en 180° en torno al enlace glicosídico y alrededor de 0° en torno al enlace aglicónico), que no existe en el compuesto natural.

Este resultado computacional fue corroborado mediante el análisis de los datos obtenidos por RMN, basado en la combinación de constantes de acoplamiento interglicosídicas, así como en la observación de NOEs inter-residuo exclusivos, que caracterizan inequívocamente los conformeros *syn*, *anti- ψ* y *anti- ϕ* . Además, la presencia de dos protones en el metileno puente permitió la observación de otros cinco NOEs que proporcionan información conformacional adicional. El ajuste entre los datos experimentales (constantes de acoplamiento y NOEs) y los calculados a partir de la distribución de población teórica confirmaron los resultados cuantitativamente.

Casi simultáneamente, y de manera independiente, Sinaý²⁷ y colaboradores analizaron el comportamiento conformacional del metil β -*C*-lactósido de forma cuantitativa mediante experimentos *off-resonance* ROESY. Estos autores llegaron a

conclusiones similares a las obtenidas en nuestro grupo de investigación, ya que los datos experimentales sólo se podían explicar con la existencia de dos familias conformacionales distintas alrededor del enlace ψ (60% *syn* y 40% *anti*). Sin embargo, estos autores no llegaron a detectar la presencia del conformero *anti- ϕ* .

La C-lactosa no es un ejemplo aislado donde se haya observado un comportamiento distinto frente al análogo natural. Así, otros estudios comparativos, como el del O/C- α -manobiósido (α -Man-(1 \rightarrow 2)-Man- α -OMe)²⁸ o el del O/C- α -Man(1 \rightarrow 1)-Gal- β -OMe²⁹, llevaron a conclusiones similares. Por tanto, aunque durante algún tiempo se asumió que los C-glicósidos tenían el mismo comportamiento conformacional que sus análogos naturales, el análisis detallado de varios C-disacáridos mediante la combinación de datos experimentales de RMN y de modelado molecular ha demostrado que esto no siempre es así.

Después del estudio de varios C-glicósidos se han podido sacar conclusiones generales acerca de su comportamiento conformacional. Al igual que la conformación de los glicósidos, la de los C-análogos viene definida por los ángulos ϕ y ψ , pero estos últimos presentan una mayor flexibilidad en torno a enlace aglicónico, ψ . En cuanto al comportamiento en torno al enlace glicosídico ϕ , se ha visto que cuando el hidroxilo de la posición 2 (OH2) se encuentra en disposición ecuatorial (análogos de tipo glucosa y galactosa) la distribución conformacional en torno a este enlace es cualitativamente parecida a la del análogo natural. No obstante, mientras que en O-glicósidos se observa casi exclusivamente la conformación *syn*, en C-glicósidos se han llegado a observar conformaciones *anti- ϕ* en análogos β ³⁰ y orientaciones *non exo*-anoméricas, en ambos análogos, α y β ³⁰. Por el contrario, cuando el OH2 adopta una orientación axial (análogos de la manosa), el comportamiento observado es muy diferente al del compuesto natural, con la presencia de una población significativa del conformero *non exo*³⁰.

Por tanto, los estudios de nuestro grupo de investigación demostraron que, la conclusión general de Kishi, apostando que los O- y los C-glicósidos tienen el mismo comportamiento conformacional, no tiene validez general y que es necesario estudiar cada caso. Además, demostramos que la combinación mecánica molecular y de parámetros de RMN (constantes de acoplamiento y NOEs) permite caracterizar el

comportamiento conformacional de este tipo de compuestos en disolución de forma cuantitativa.

En esta memoria, con objeto de extender el estudio comparativo de las propiedades conformacionales de los *C*- y *O*-glicósidos a otros disacáridos, se ha llevado a cabo un análisis conformacional del disacárido β -Gal-(1 \rightarrow 3)Glc- β -OMe y su *C*-análogo siguiendo una metodología similar a la descrita. Este *C*-disacárido es de gran interés biológico por ser el *C*-análogo de uno de los disacáridos constituyentes del sistema inmunológico Lewis. Además, este estudio servirá de punto de partida para el análisis posterior del Lewis-A y su *C*-análogo (capítulo 4).

2. Análisis conformacional del disacárido β -Gal(1 \rightarrow 3) β -Glc-OMe

Para referirnos al disacárido β -Gal(1 \rightarrow 3) β -Glc-OMe (**1**), se utilizará la nomenclatura habitual en química de carbohidratos, de tal forma que Gal corresponde al residuo de *D*-galactosa y Glc al de *D*-glucosa que, en este caso, está bloqueada en su forma anomérica β por un grupo metilo (**Figura 2.4**).

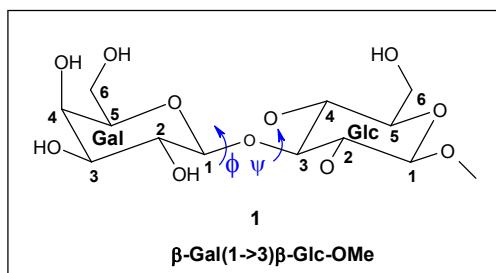


Figura 2.4: Estructura del disacárido β -Gal(1 \rightarrow 3) β -Glc-OMe, junto con la numeración y la designación de los ángulos diedros que serán utilizados en esta memoria.

El protocolo utilizado para deducir el comportamiento conformacional de este tipo de compuestos involucra el uso de cálculos de mecánica molecular para la determinación del mapa (ϕ, ψ) de energía conformacional, seguido del cálculo de la distribución de Boltzmann, a partir del cual se calculan los NOEs esperados para cada pareja de protones, que se comparan con los NOEs experimentales para determinar si la distribución teórica se ajusta con la real en disolución.

2.1. Cálculos de Mecánica Molecular:

Para la determinación de los mapas conformacionales, se realizó una búsqueda sistemática, eligiendo MM3* como campo de fuerzas y una constante dieléctrica de 80 debyes (para simular el disolvente, H₂O). Estudios anteriores habían demostrado que esta elección permitía conseguir un buen acuerdo entre datos experimentales y teóricos en otros disacáridos³¹. Los ángulos diedros fueron definidos como ϕ (Gal1-GalC1-O-GlcC3) y ψ (GalC1-O-GlcC3-Glc3). Se calcularon cuatro mapas relajados teniendo en cuenta las distintas orientaciones posibles del grupo hidroximetilo (*gt* y *tg* para el anillo de galactosa y *gg* y *gt* para el de glucosa). A partir de estos mapas relajados, se calcularon las superficies adiabáticas, así como la distribución de población, de acuerdo con la función de Boltzmann.

Los cálculos indicaron que casi la totalidad de la población (alrededor del 98%) se concentra en la región central *syn* comprendida entre 0° y 100° para ϕ y -60° y 90° para ψ (**Figura 2.5**), representada por el mínimo A (**Figura 2.6, Tabla 2.2**). Las otras dos regiones de baja energía apenas concentran un 1% de la población cada una de ellas. Por un lado, la región *anti- ψ* , representada por el mínimo B, presenta valores de ϕ similares a los de la región *syn*, pero muestra valores de ψ centrados en torno a 180°. Por otro, la región *anti- ϕ* , representada por el mínimo C, presenta valores de ϕ centrados en torno a 180° y valores de ψ cercanos a los de la región *syn*.

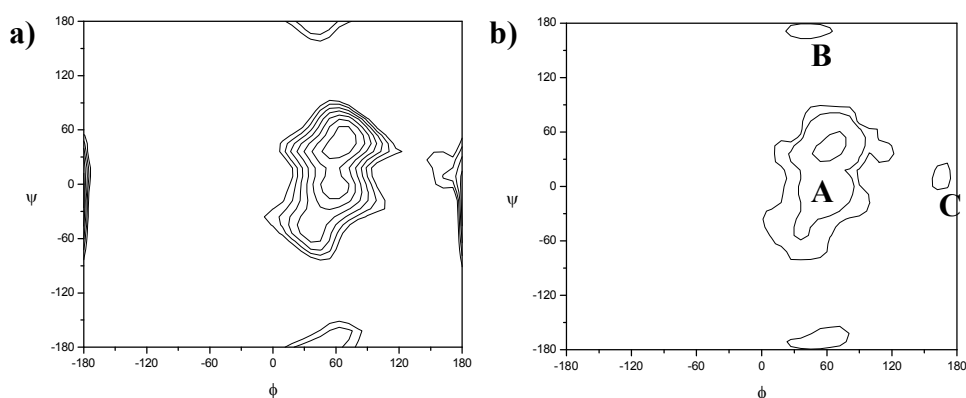


Figura 2.5: Mapa adiabático (a) y de probabilidad (b) para **1**. Los contornos del mapa energético se representan cada 0.5 kcal/mol. En la distribución de probabilidad, las curvas de nivel se representan para poblaciones de 0.1%, 1% y 10%.

		β -Gal(1 \rightarrow 3) β -Glc-OMe (1)		
		<i>syn</i> (A)	<i>anti-ψ</i> (B)	<i>anti-ϕ</i> (C)
Gal(1 \rightarrow 3)Glc	ϕ	57	41	165
	ψ	25	168	16
Población		98	1	1

Tabla 2.2: Ángulos diedros y poblaciones (%) para los mínimos conformacionales de este disacárido de acuerdo con el campo de fuerzas MM3*.

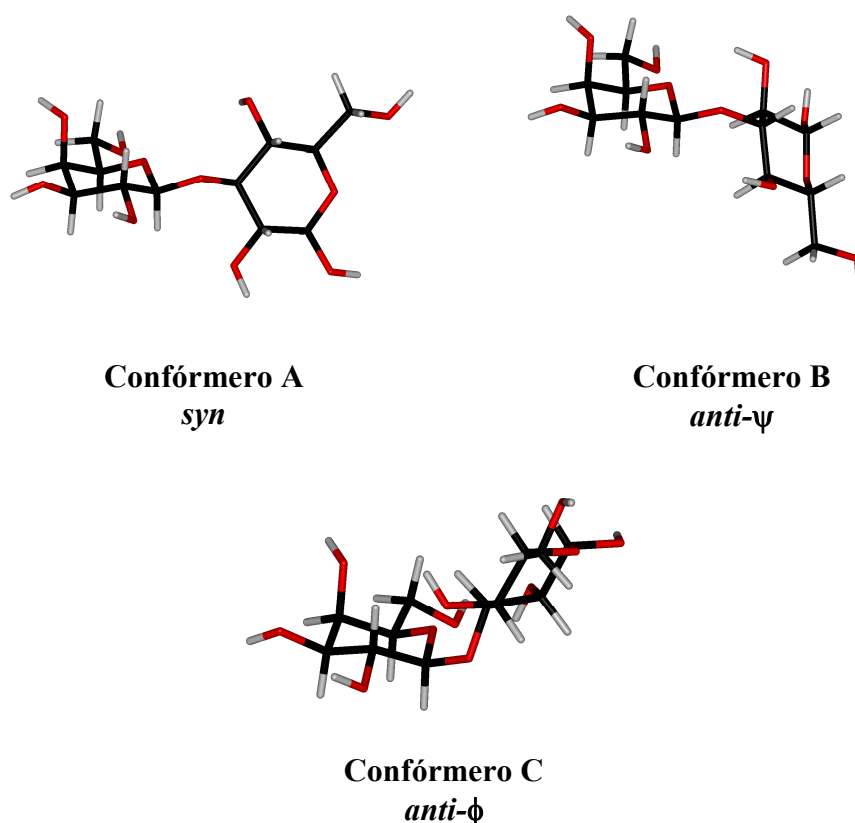


Figura 2.6: Geometría de los mínimos conformacionales del *O*-glicósido, *syn* (A), *anti- ψ* (B) y *anti- ϕ* (C).

Estos resultados concuerdan con los estudios computacionales realizados anteriormente para uniones $\beta(1\rightarrow3)$, como el estudio del β -laminarabiósido (β -Glc(1 \rightarrow 3) β Glc)³², o el disacárido β -Gal(1 \rightarrow 3)GlcNAc³³⁻³⁴. Más recientemente, también en el caso del β -laminarabiósido, se ha investigado, a través de simulaciones de dinámica molecular en agua, la movilidad conformacional alrededor del enlace $\beta(1\rightarrow3)$. Los ángulos diedros exhibieron un bajo grado de flexibilidad, manteniendo la trayectoria en la zona central de menor energía³⁵. Los estudios en disolución de este compuesto mostraron que los ángulos diedros ϕ y ψ , derivados de las constantes de acoplamiento protón-carbono $^3J_{\text{Gal1-GlcC3}}$ y $^3J_{\text{GalC1-Glc3}}$, eran consistentes con el predominio de la conformación *syn*³⁶. Por otro lado, el análisis cristalográfico de diferentes oligosacáridos conteniendo la unión $\beta(1\rightarrow3)$, puso de manifiesto que esta unión sólo adopta la conformación *syn* en estado sólido, en total acuerdo con los datos teóricos³⁷.

2.2. Estudio mediante RMN:

Siguiendo el protocolo descrito en el capítulo 1, se asignaron las resonancias de todos los protones mediante la combinación de experimentos de RMN (**Figura 2.7**), 1D-TOCSY y 2D-HSQC a 500 MHz. En la **Tabla 2.3** se recogen los desplazamientos químicos de ^1H RMN.

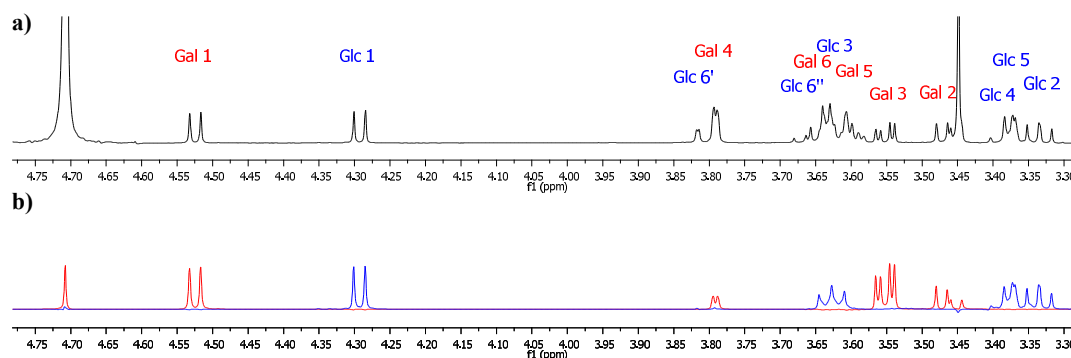


Figura 2.7: Espectro de protón (a) y espectros 1D-TOCSY (b) en los que se obtuvo el subespectro de la unidad de galactosa (en rojo) y de la de glucosa (en azul) por excitación selectiva de cada uno de los protones anoméricos. Los espectros se adquirieron a 25°C en D_2O y a 500 MHz.

1	Gal	Glc
H1	4.52	4.28
H2	3.46	3.33
H3	3.55	3.63
H4	3.79	3.37
H5	3.59	3.36
H6a	3.63	3.65
H6b	3.63	3.79
Me	-	3.44

Tabla 2.3: Desplazamientos químicos de ^1H RMN (δ , ppm) en D_2O a 25°C y a 500 MHz.

Para el análisis de los NOEs, se realizaron experimentos 1D-NOESY y 1D-ROESY a varios tiempos de mezcla (**Figura 2.8**), en los que el protón anomérico de la galactosa (Gal1) fue excitado selectivamente y experimentos 2D-NOESY. Además, como una de las señales clave, Glc3, aparecía solapada con otras resonancias, se realizó un experimento 2D-HSQC-NOESY para separar los NOEs en la dimensión de carbono, ya que esta presenta una mayor dispersión. En la **Figura**

2.9 se puede observar el espectro 2D-HSQC-NOESY, en el que se han indicado los NOEs interglucosídicos claves para el estudio conformacional.

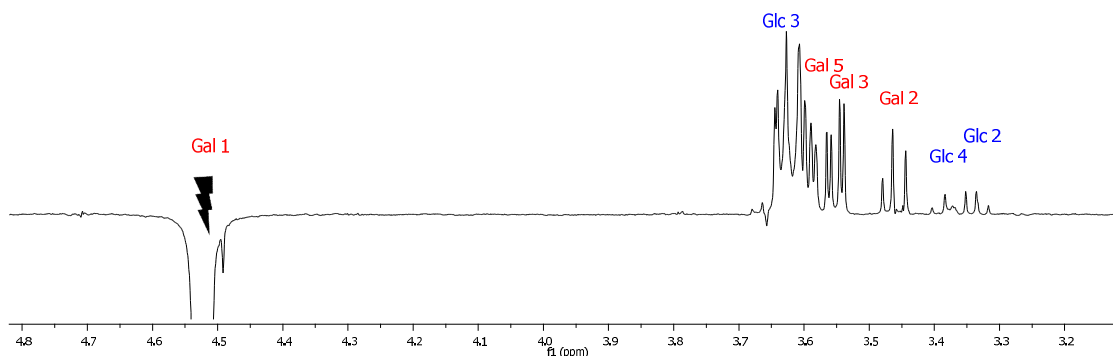


Figura 2.8: Espectro 1D-NOESY con un tiempo de mezcla de 500 ms, donde se invirtió el protón Gal1. El experimento se adquirió a 25°C en D₂O y a 500 MHz.

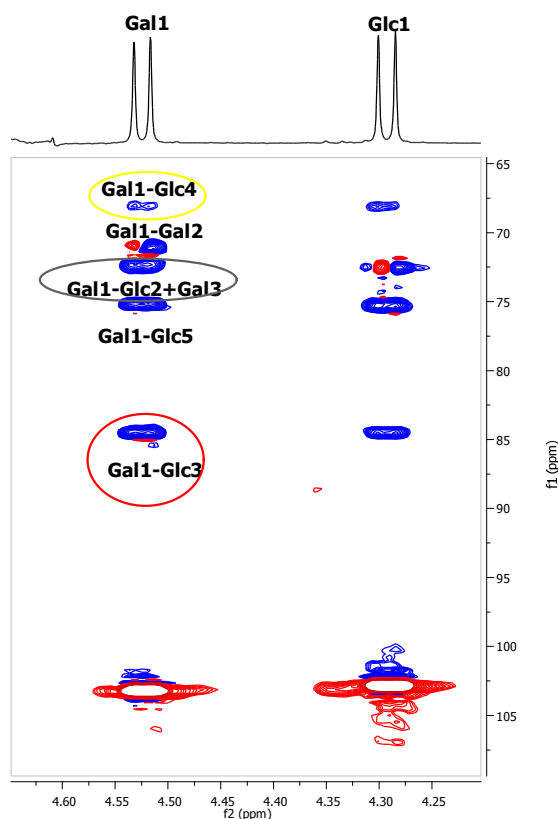


Figura 2.9: Expansión del espectro 2D-HSQC-NOESY a 500 MHz, en la que se observan los NOEs más importantes clasificados mediante código de color en intenso (rojo) y débil (amarillo). El NOE Gal1-Glc2 no se pudo clasificar por su solapamiento con Gal1-Gal3 (en gris).

Cada una de las familias conformacionales detectadas en los cálculos de mecánica molecular viene caracterizada por un NOE interglucosídico, de acuerdo a las distancias entre protones recogidas en la **Tabla 2.4**. El hecho de que el NOE

Gal1-Glc3, característico del conformero *syn*, sea el más intenso en todos los espectros NOESY indica que, al menos cualitativamente, esta familia conformacional es la predominante en disolución, lo que está de acuerdo con los resultados computacionales. Por otro lado, dado que el NOE exclusivo del conformero *anti- ϕ* , Gal2-Glc3, no se detecta en ninguno de los experimentos NOESY, se puede afirmar que la región descrita por este conformero no se encuentra poblada de manera significativa. Además del NOE intenso Gal1-Glc3, se observan otros dos NOEs débiles, que son informativos desde el punto de vista conformacional. Se trata de los NOEs Gal1-Glc2 y Gal1-Glc4, este último exclusivo del conformero *anti- ψ* , lo que indicaría la presencia de esta familia conformacional en disolución, aunque de forma minoritaria.

		<i>syn</i> (A)	<i>anti-ψ</i> (B)	<i>anti-ϕ</i> (C)	Distribución
Gal1	Glc2	3.3	2.6	4.3	3.1
	Glc4	4.5	2.0	4.4	3.8
	Glc3	2.8	3.7	3.7	2.7
Gal2	Glc3	4.3	4.5	2.0	4.1

Tabla 2.4: Distancias interprotónicas (Å) para cada uno de los conformeros. Las distancias menores de 3 Å están señaladas en azul, indicando los NOEs característicos de los conformeros *anti- ϕ* y *anti- ψ* . Las distancias promedio de la distribución corresponden a una distribución $\langle r^{-6} \rangle^{-1/6}$.

En la **Tabla 2.5** se recogen los valores de las intensidades de NOEs teóricas para cada uno de los conformeros, y para la distribución de Boltzmann, utilizando la aproximación de la matriz de relajación completa³⁸⁻³⁹. Además, se muestran los valores medidos en el experimento 1D-NOESY. De estos datos se puede deducir que si el conformero *syn* fuera el único poblado en disolución, además de un NOE intenso entre los protones Gal1 y Glc3, se debería observar el NOE Gal1-Glc2 con una intensidad en torno al 1%, mientras que el NOE Gal1-Glc4 no sería prácticamente detectable (intensidad del 0.1%, en el límite de detección experimental). Sin embargo, estos dos NOEs presentan una intensidad parecida en todos los experimentos NOESY realizados a varios tiempos de mezcla, lo que sólo puede justificarse con una población del conformero *anti- ψ* mayor que la del 1%, que se predice mediante los cálculos de mecánica molecular.

Para tratar de estimar la proporción *syn:anti- ψ* que mejor se ajustaría a los resultados experimentales, se determinaron los NOEs teóricos para distintas

proporciones de los conformeros *syn* y *anti- ψ* , tomándose los mínimos A y B como los conformeros representativos de cada familia conformacional. De esta manera, se determinó que, con una proporción 95:5 de los conformeros A y B, la intensidad de los NOEs Gal1-Glc2 y Gal1-Glc4 sería muy parecida, manteniéndose un buen ajuste entre los valores experimentales y calculados.

		NOEs Teóricos				NOEs Experimentales
		<i>syn</i> (A)	<i>anti-ψ</i> (B)	<i>Anti-ϕ</i> (C)	Distribución	
Gal1	Gal3	2.5	2.8	2.6	2.6	2.5
	Glc2	1.0	2.3	0.2	0.9	0.4
	Glc3	2.5	0.3	0.3	2.4	4.5*
	Glc4	0.1	9.5	0.1	0.3	0.4
Gal2	Glc3	0.1	0.1	6.7	0.2	-

Tabla 2.5: Intensidades de NOE (%) calculadas mediante la aproximación de la matriz de relajación completa, junto a los valores experimentales obtenidos en experimentos 1D-NOESY con un tiempo de mezcla de 500 ms. (*Debido al solapamiento de Glc3 y Gal5, el NOE Gal1-Glc3 se estimó como la diferencia entre la integral correspondiente al NOE Gal1-Glc3+Gal5 y la intensidad teórica del NOE intra-residuo Gal1-Gal3 que no depende de la conformación interglucosídica y que se asume de intensidad similar al Gal1 Gal5 ya que las distancias intra-residuo H1-H3 y H1-H5 son semejantes.)

2.3. Conclusiones:

La combinación de los estudios de RMN y de mecánica molecular indica que la distribución conformacional del *O*-disacárido **1** está dominada por la familia conformacional *syn*, como se demostró para otros disacáridos uniones $\beta(1\rightarrow3)$. Sin embargo, un análisis detallado de la intensidad de los NOEs interglucosídicos, pone de manifiesto la existencia de una segunda familia conformacional *anti- ψ* , con una población aproximada del 5%.

La existencia de la familia conformacional *anti- ψ* para este disacárido fue propuesta anteriormente por Dabrowski y colaboradores³² mediante la detección de un enlace de hidrógeno intramolecular que sólo puede darse en esta conformación. Sin embargo, estos autores realizaron sus estudios en un disolvente orgánico, DMSO, por lo que nuestros resultados constituyen la primera evidencia experimental de que al menos una pequeña, pero detectable, población del conformero *anti- ψ* coexiste con el conformero mayoritario *syn* en disolución acuosa.

3. Análisis conformacional del C-disacárido β -C-Gal(1 \rightarrow 3)- β -Glc-OMe:

Este compuesto (**2**) es análogo al compuesto **1**, pero el oxígeno glicosídico entre las dos unidades de monosacáridos ha sido sustituido por un grupo metileno. También en este caso, el C-disacárido se encuentra bloqueado en su forma anomérica β por la presencia de un grupo metilo (**Figura 2.10**).

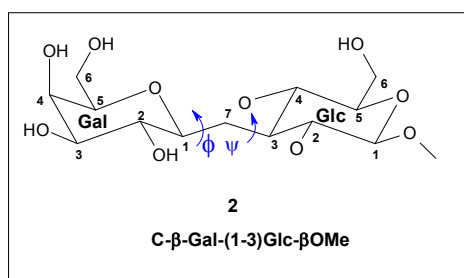


Figura 2.10: Estructura del C-disacárido β -C-Gal(1 \rightarrow 3) β Glc-OMe junto con la numeración y la designación de los ángulos diedros que serán utilizados en esta tesis.

3.1. Cálculos de Mecánica Molecular:

De la misma manera que en el O-disacárido, se realizaron cálculos de mecánica molecular para la obtención de los mapas relajados y adiabáticos, y a partir de ellos se determinó la distribución de probabilidad, de acuerdo con la ecuación de Boltzmann. Por tanto, se realizó una búsqueda sistemática con el campo de fuerzas MM3* y una constante dieléctrica de 80 debyes, calculando todos los conformeros posibles en torno a los ángulos diedros ϕ y ψ , definidos de manera análoga al O-disacárido, como Gal1-GalC1-C7-GlcC3 y GalC1-C7-GlcC3-Glc3, respectivamente. Como en el caso anterior, se tuvieron en cuenta las disposiciones del grupo hidroximetilo *gt* y *tg* para la galactosa y *gg* y *gt* para la glucosa.

Mientras que en el O-disacárido, los cálculos indicaron que casi la totalidad de la población se concentraba en la zona central del mapa, en este caso, la población aparece muy repartida y se predice la existencia de cinco las familias conformacionales en disolución (**Figura 2.11**). De nuevo, la región central es la zona más poblada, con un 72% de la población, aunque la región es algo más extensa que en el O-disacárido, estando delimitada por valores de ϕ entre 0 y 100° y valores de ψ entre -90 y 90°. Esta región está realmente dividida en dos zonas, representadas por los mínimos **A** y **D** (**Figura 2.12**) con una población del 60 y el 12%

respectivamente. Los dos mínimos presentan valores del ángulo ϕ parecidos, pero se diferencian en el valor del ángulo ψ y, en consecuencia, se designaron como *syn*- ψ (+) y *syn*- ψ (-). Por otro lado, la diferencia más importante con respecto a **1** es que la región *anti*- ψ , representada por el mínimo **B**, se encuentra significativamente poblada de acuerdo con los datos de MM3*, concentrando el 24% de la población; mientras que ésta sólo era en torno al 1% en el *O*-disacárido. Finalmente, las otras dos familias conformacionales, *anti*- ϕ (mínimo C) y *non exo* (mínimo E), apenas están pobladas, concentrando el 1 y 3% de la población, respectivamente (Tabla 2.6).

		β -C-Gal(1 \rightarrow 3)- β Glc-OMe (2)				
		<i>syn</i> - ψ (+) (A)	<i>anti</i> - ψ (B)	<i>anti</i> - ϕ (C)	<i>syn</i> - ψ (-) (D)	<i>non exo</i> (E)
C-Gal(1 \rightarrow 3)Glc	ϕ	65	45	-171	38	-57
	ψ	56	176	53	-59	-54
Población		60	24	1	12	3

Tabla 2.6: Ángulos diedros y poblaciones (%) para los mínimos conformacionales del C-glicósido. Los conformeros se han denotado como **A**, **B**, **C**, **D** y **E**.

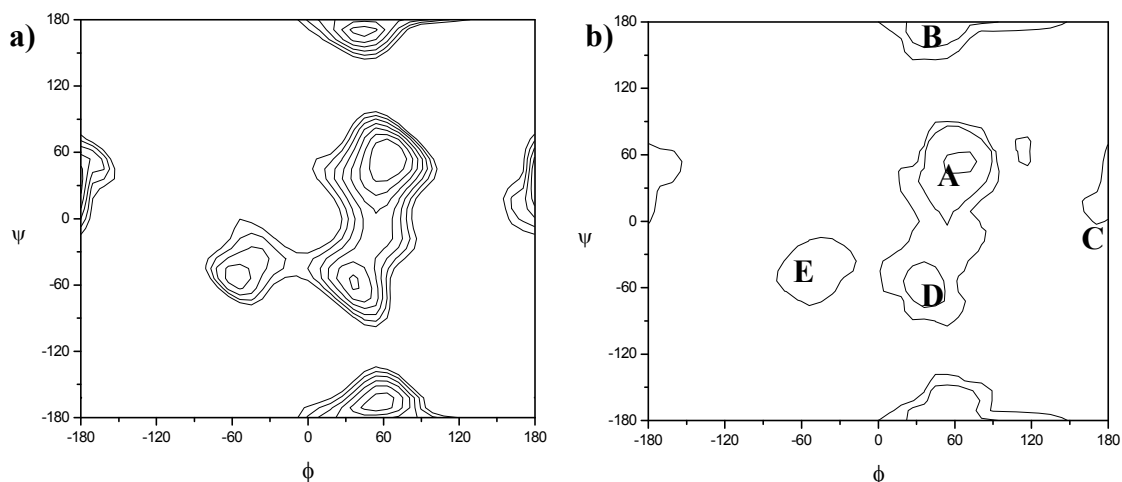
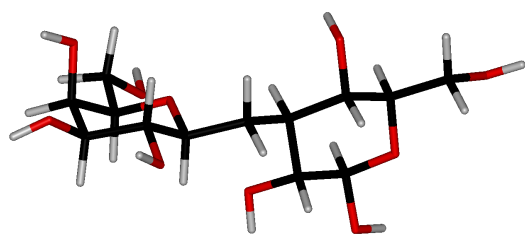
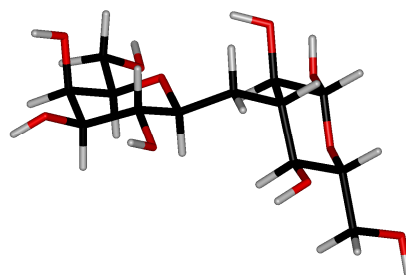


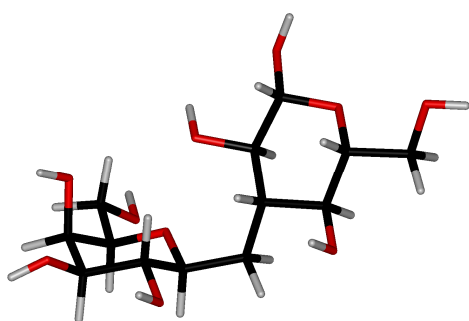
Figura 2.11: Mapa adiabático (a) y de probabilidad (b) para **2**. Los contornos del mapa energético se representan cada 0.5 kcal/mol. En la distribución de probabilidad, las curvas de nivel se representan para poblaciones de 0.1%, 1% y 10%.



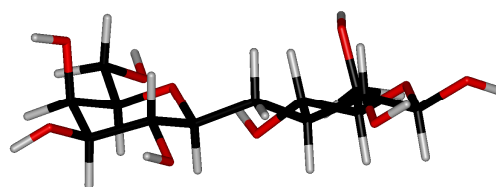
Confórmero A
syn- ψ (+)



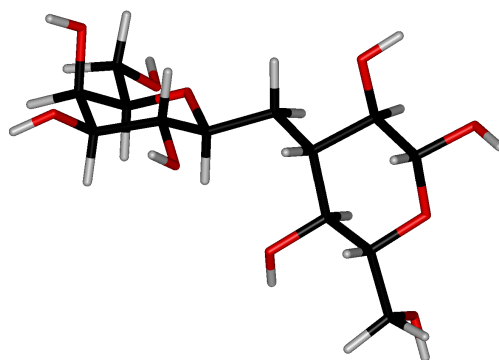
Confórmero B
anti- ψ



Confórmero C
anti- ϕ



Confórmero D
syn- ψ (-)



Confórmero E
non exo

Figura 2.12: Geometría de los mínimos conformaciones del C-glicósido, A-E.

3.2. Estudio mediante RMN:

La asignación de los desplazamientos químicos de todos los protones de **2** se obtuvo mediante experimentos de ^1H , COSY y HSQC. La asignación de los protones proquirales 7_{pR} y 7_{pS} del puente metilénico se realizó usando una combinación de constantes protón-protón y NOEs. Los desplazamientos químicos de protón en D_2O y MeOD están recogidos en la **Tabla 2.7**.

2	D_2O		MeOD	
	Gal	Glc	Gal	Glc
H1	3.45	4.33	3.40	4.15
H2	3.42	3.31	3.42	3.15
H3	3.60	1.8	3.44	1.69
H4	3.94	3.47	3.84	3.21
H5	3.62	3.97	3.48	3.31
H6a	3.72	3.69	3.73	3.87
H6b	3.69	3.55	3.67	3.64
OMe		3.55		3.53
7_{pR}	2.23		2.34	
7_{pS}	1.79		1.76	

Tabla 2.7: Desplazamientos químicos de protón (en ppm) en D_2O y MeOD a 25°C y 500 MHz.

La presencia del puente metilénico en el C-disacárido es una gran ventaja a la hora de realizar un estudio conformacional de este compuesto, ya que las cuatro constantes de acoplamiento a través del enlace pseudo-glicosídico proporcionan información de los ángulos diedros ϕ y ψ , a través de la ecuación de Karplus-Altona⁴⁰. Para la determinación de valores exactos de estas constantes de acoplamiento, fue necesario llevar a cabo un análisis de segundo orden. El mejor ajuste entre el espectro experimental y el teórico se muestra en la **Figura 2.13**. Los valores determinados experimentalmente se compararon con los valores calculados para los mínimos A-E, así como con los valores promedio de la distribución teórica, obtenida mediante mecánica molecular (**Tabla 2.8**).

El valor pequeño de la constante de acoplamiento $^3J_{\text{Gal1-H7pS}}$ (1.3 Hz en D_2O) indica que el conformero *non exo* (E) apenas está poblado en disolución ya que, en esta geometría, la disposición de los protones Gal1 y H7pS es *anti*-periplanar, por lo que la constante de acoplamiento sería mucho mayor si este conformero estuviera poblado de manera significativa. De la misma manera, el valor de la constante de

acoplamiento $^3J_{\text{Gal1-H7pR}}$ (9.8 Hz en D₂O), aunque no tan extrema como en el caso anterior descarta también la presencia de una población apreciable de la familia conformacional *anti- ϕ* (C), ya que, en este conformero, los protones Gal1 y H7pS quedan dispuestos en orientación *gauche*, lo que se traduciría en un valor de $^3J_{\text{Gal1-H7pR}}$ mucho más pequeño. Estos datos experimentales concuerdan, al menos cualitativamente, con los resultados obtenidos mediante los cálculos de mecánica molecular, ya que ambas regiones aparecen muy poco pobladas en el mapa de distribución de población predicho por MM3*.

	<i>syn-ψ</i> (+) (A)	<i>anti-ψ</i> (B)	<i>anti-ϕ</i> (C)	<i>syn-ψ</i> (-) (D)	<i>non exo</i> (E)	Distribución teórica	Exp. D ₂ O	Exp. MeOD
Gal1-H7pS	2,3	1,0	3,3	1,0	11,5	2,3	1,3	1,6
Gal1-H7pR	11,7	10,7	3,4	9,9	3,3	10,6	9,8	8,9
Glc3-H7pS	12,4	2,8	12,3	2,9	2,1	8,3	5,5	5,4
Glc3-H7pR	2,5	4,2	2,1	12,3	12,3	3,9	4,1	4,8

Tabla 2.8: Constantes de acoplamiento (J, Hz) a través del enlace pseudo-glicosídico para los mínimos A-E. Se muestran los datos esperados para la distribución teórica propuesta por MM3*, junto con los valores experimentales en D₂O y MeOD.

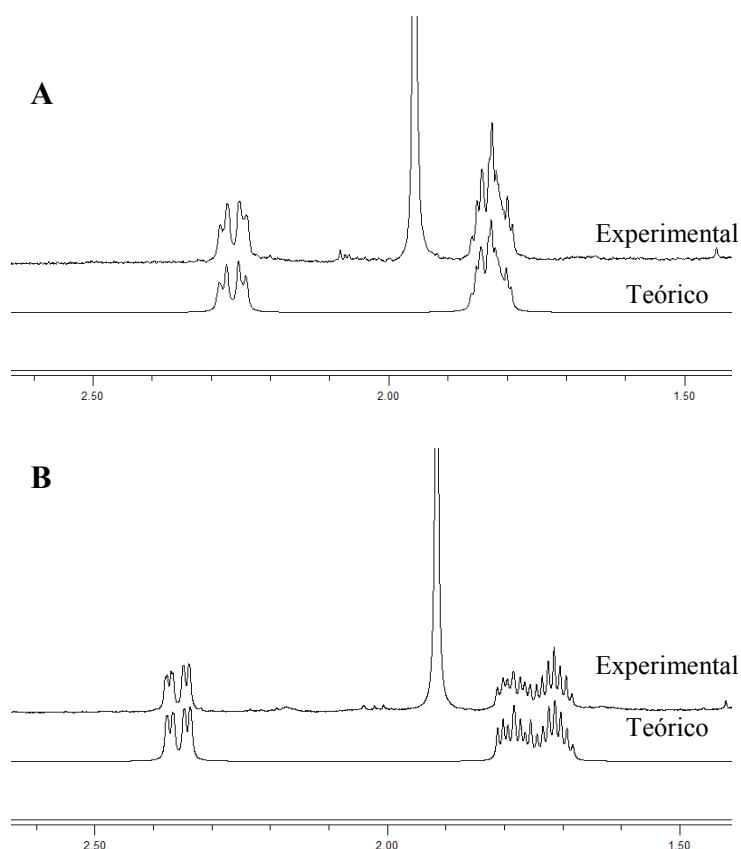


Figura 2.13: Espectros simulado y experimental tanto en D₂O (A) como en MeOD (B) para la determinación de las constantes de acoplamiento protón-protón a 500 MHz.

Por tanto, los valores extremos obtenidos para las constantes de acoplamiento, entre Gal1 y los dos protones metilénicos, indican que el C-disacárido adopta predominantemente la conformación *exo*-anomérica (alrededor de 60°) en torno al enlace pseudo-glicosídico, como ocurre en los correspondientes O-disacáridos. Por el contrario, los valores intermedios obtenidos a través del enlace pseudo-aglicónico ($^3J_{H3-H7pS}$ and $^3J_{H3-H7pR}$) reflejan la existencia de un equilibrio conformacional en torno a este enlace, que conduciría a constantes de acoplamiento promedio. Este dato está de acuerdo con la coexistencia de las conformaciones *syn*- ψ y *anti*- ψ que predicen los estudios computacionales. Aunque el ajuste entre los valores obtenidos de la distribución y los valores experimentales es aceptable, existe una diferencia notable entre el valor calculado y experimental para la constante $J_{H3-H7pS}$, cuyo valor es mayor en la distribución teórica. Este resultado sugiere que los cálculos sobrestiman la población del conformero *syn*- ψ (+) ya que, en esta conformación, estos protones están dispuestos en una orientación *anti*-periplanar, lo que daría lugar a una constante de acoplamiento más grande. Por otro lado, para mejorar el ajuste de esta constante sin empeorar el del resto, sería necesario que la población del mínimo *anti*- ψ fuera mayor.

Además, se realizaron experimentos 1D-NOESY, 2D-NOESY y 2D-ROESY para complementar la información obtenida mediante constantes de acoplamiento. Desafortunadamente, el análisis de NOEs en D₂O no fue posible debido al solapamiento de varias señales claves: Glc2 y Glc4 aparecen con el mismo desplazamiento químico y Glc3 y H7pR solapan parcialmente. Debido a la escasa cantidad de muestra disponible, tampoco fue posible realizar un experimento 2D-HSQC-NOESY, que hubiera ayudado a separar los NOEs en la dimensión de carbono, como se hizo en el O-disacárido. Para solucionar este problema se registraron espectros a distintas temperaturas, pero, desafortunadamente, no se consiguió mejorar la situación. Finalmente, se optó por llevar a cabo el estudio en metanol-d₄ (MeOD), donde se observó una mayor dispersión. A la hora de cambiar de disolvente hay que tener en cuenta que se pueden producir cambios conformacionales importantes. Sin embargo, el hecho de que las constantes de acoplamiento sean similares a las obtenidas en D₂O sugiere que el comportamiento conformacional del C-disacárido es similar en ambos disolventes, por lo que los datos de NOE en MeOD serían relevantes.

Las distancias interglucosídicas más informativas, desde el punto de vista conformacional, se recogen en la **Tabla 2.9** para los mínimos A, B y D (por simplicidad los mínimos C y E no se han incluido ya que, mediante el análisis de las constantes de acoplamiento, se concluyó que no estaban poblados o su población era pequeña). A partir de estas distancias, se puede deducir que hay dos NOEs, Gal1-Glc3 y Gal1-Glc4, que caracterizan las orientaciones *syn*- ψ y *anti*- ψ , respectivamente. La observación de ambos NOEs en el espectro (**Figura 2.14**) confirmó lo que ya se había deducido mediante constantes de acoplamiento protón-protón: la coexistencia de ambos conformeros en disolución.

		<i>syn</i> - ψ (+) (A)	<i>anti</i> - ψ (B)	<i>syn</i> - ψ (-) (D)
Gal1	Glc2	3.0	2.8	4.7
	Glc4	4.5	2.2	3.8
	Glc3	3.3	3.8	2.6
Glc2	Hp7R	3.8	3.3	2.4
	Hp7S	2.5	3.9	2.9
Glc3	Hp7R	2.5	2.4	3.1
	Hp7S	3.1	2.5	2.5
Glc4	Hp7R	2.9	3.8	2.6
	Hp7S	2.5	3.2	3.8

Tabla 2.9: Distancias interprotónicas (Å) para cada uno de los conformeros.

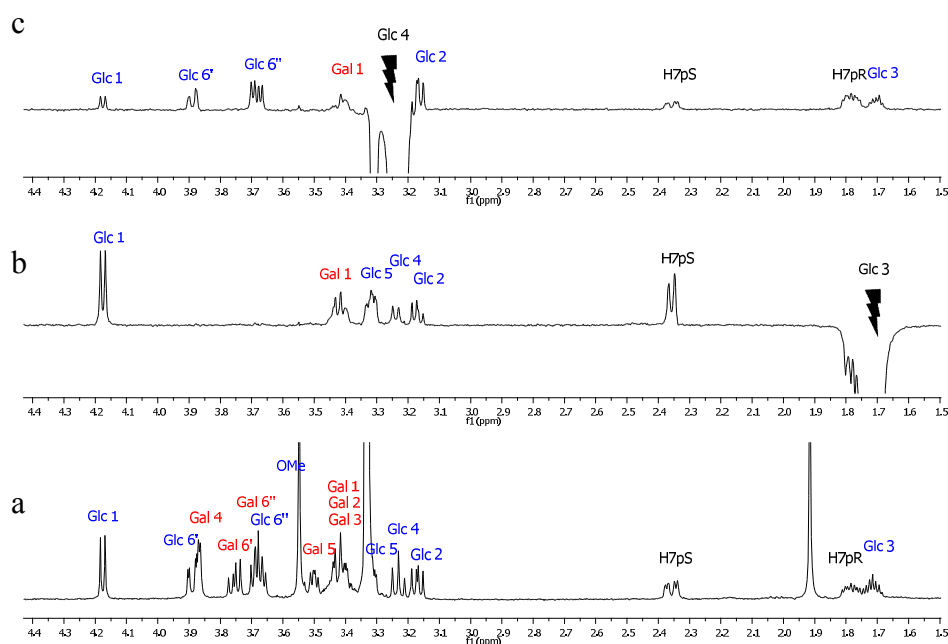


Figura 2.14: Espectros de protón (a) y 1D-NOESY en los que se excitaron las señales Glc3(b) y Glc4 (c). El tiempo de mezcla de los experimentos 1D-NOESY fue de 500 ms en todos los casos a 500MHz.

La mayor intensidad del NOE Gal1-Glc3 con respecto al NOE Gal1-Glc4 indicaría, al menos cualitativamente, una mayor abundancia del conformero *syn-ψ* en el equilibrio conformacional. Para realizar un estudio más cuantitativo, se recurrió, como en **1**, a la matriz de relajación completa, mediante el programa NOEPROM, desarrollado en nuestro grupo de trabajo, para calcular los NOEs esperados para la distribución teórica de MM3* (**Tabla 2.10**).

		NOEs Teóricos (%)				NOEs Experimentales
		<i>synψ</i> (+) (A)	<i>anti-ψ</i> (B)	<i>syn-ψ</i> (-) (D)	Distribución	%NOE
Glc2	H7pR	-0.1	0.4	2.8	0.4	0.4
	H7pS	2.3	0.1	0.6	1.6	1.5
	Gal1	1.1	1.4	0.0	1.0	0.3
Glc3	H7pR	2.2	2.9	0.4	2.2	Solapamiento
	H7pS	0.4	2.5	2.6	1.1	2.4
	Gal1	0.7	0.2	3.0	0.8	2.5
Glc4	H7pR	0.6	0.1	2.0	0.7	1.2
	H7pS	2.4	0.6	0.0	1.7	0.5
	Gal1	0.0	7.1	0.3	1.7	1.1

Tabla 2.10: NOEs teóricos (%) de los conformeros A, B y D y de la distribución de Boltzmann Estimada mediante los cálculos MM3*, frente a los NOEs experimentales medidos en los espectros 1D-NOESY, con un tiempo de mezcla de 500 ms a 500 MHz. El protón de la primera columna es el que se excita selectivamente en los experimentos 1D-NOESY.

La comparación de los valores de los NOEs experimentales con los calculados indica que el ajuste es razonable, aunque existen diferencias claras entre algunos de los valores. Las mayores discrepancias se encuentran en los NOEs Glc4-H7ps y Glc3-Gal1. Así, la intensidad experimental de los NOEs Glc4-H7ps y Glc2-Gal1 es notablemente menor que la esperada para la distribución. Ya que estos son NOEs esperados para la conformación *syn-ψ*(+), se llega a la misma conclusión a partir de las constantes de acoplamiento que los cálculos MM3* sobrestiman la población de la familia conformacional *syn-ψ*(+). Por otro lado, la intensidad experimental del NOE Glc3-Gal1 es mayor que la esperada, lo que indicaría que la familia conformacional *syn-ψ*(-), que es aquella en la que este NOE es más intenso, está realmente más poblada que lo que se predice mediante los cálculos teóricos con el campo de fuerzas MM3*. En el capítulo 5 de esta tesis se vuelve a abordar el análisis

conformacional de este compuesto mediante un método alternativo (JAMFIS) que busca la población de cada conformero que produce el mejor ajuste entre constantes de acoplamiento experimentales y teóricas, concluyéndose de nuevo que la población del conformero *anti*- ψ es mucho mayor que la determinada por MM3*, mientras que la del conformero *syn*- ψ (+) es mucho menor.

4. Conclusiones del análisis comparativo O/C-β-Gal(1→3)β-Glc-OMe:

La combinación de estudios de RMN y de mecánica molecular indica que la distribución conformacional del C-disacárido **2** se concentra en dos zonas distintas del mapa conformacional. La región central, *syn*-ψ, dividida en dos zonas, *syn*-ψ(+) y *syn*-ψ(-), representadas por los mínimos A y D. La región *anti*-ψ está representada por el mínimo B, con aproximadamente un 30% de la población. Todas las conformaciones presentan valores similares del ángulo φ, *exo*-anomérico, pero se diferencian en el valor del ángulo ψ, que adopta una disposición *gauche*(+) o *gauche*(-) en las conformaciones *syn*-ψ, y *anti*-periplanar, en la conformación *anti*-ψ.

Por tanto, este resultado pone de manifiesto que las propiedades conformacionales del O-disacárido (**1**) y del C-disacárido (**2**) son muy similares en torno al enlace glicosídico, adoptando preferentemente la disposición *exo*-anomérica, tal como se ha observado para otros C-disacáridos a pesar de la ausencia de estabilización por efectos estéreo-electrónicos. Sin embargo, en torno al enlace aglicónico, sí se observan algunas diferencias significativas. Mientras que en el O-disacárido la familia *syn*-ψ(-) es muy predominante y tan sólo se observa un 5% de la familia *anti*-ψ, en el C-análogo la población de esta segunda familia aumenta hasta un 24% según los cálculos de MM3* que puede considerarse un límite inferior, ya que los datos de RMN indican una población mayor. Por tanto, esta conclusión es semejante a la que se obtuvo para la lactosa, donde la conformación *syn* también era muy predominante en el compuesto natural (aproximadamente 90:10 *syn:anti*), mientras que las dos familias tenían poblaciones parecidas en el correspondiente C-análogo. Es decir, el cambio de población era más acusado en ese caso. Nuestro estudio revela que el C-disacárido (**2**) presenta una mayor flexibilidad que el correspondiente O-disacárido (**1**), lo que implicaría una mayor pérdida de entropía en el reconocimiento de **2** por parte de una proteína, pudiendo ser una limitación a la hora de usar estos análogos como inhibidores glicosídicos. No obstante, se ha demostrado que los C-glicósidos son excelentes compuestos para sondear el sitio activo de lectinas y glicosidasas⁴¹.

5. Parte Experimental:

5.1. Materiales:

El compuesto **1** y el derivado peracetilado de **2** fueron obtenidos mediante protocolos de síntesis orgánica⁴²⁻⁴³. Para la obtención del compuesto **2**, este derivado fue tratado con un exceso de metóxido sódico en metanol, a 25°C, durante una hora. Posteriormente, la disolución obtenida se agitó con Amberlyst durante 30 minutos, se filtró y se evaporó el disolvente.

5.2. Mecánica Molecular:

Las estructuras iniciales para cada uno de los mapas relajados se construyeron con el programa MAESTRO V7.0. Se tuvieron en cuenta orientaciones *tg* y *gt* para la unidad de galactosa y *gg* y *gt* para la de glucosa, ya que se ha visto que son mucho más estables que la orientación *gg* o *tg* respectivamente. Para ambos compuestos, **1** y **2**, se realizó una búsqueda sistemática mediante la rotación de los enlaces glicosídico (ϕ) y aglicónico (ψ) en pasos de 18°. Se utilizó para ello el módulo Macromodel v 8.0, seleccionando MM3* como campo de fuerzas y una constante dieléctrica de 80 debyes para simular el medio acuoso. Las estructuras generadas fueron optimizadas mediante el uso de 500 iteraciones de gradientes conjugados hasta alcanzar la convergencia. De esta manera, se calcularon 400 conformeros para construir cada mapa relajado (ϕ, ψ). El mapa de superficie adiabático se obtuvo a partir de los cuatro mapas relajados, eligiendo para cada punto (ϕ, ψ) la estructura de mínima energía. El mapa de distribución poblacional se calculó de acuerdo a la ecuación de Boltzmann a 25°C.

5.3. Cálculos de constantes de acoplamiento y NOEs:

Las constantes de acoplamiento fueron calculadas para cada conformación del compuesto **2** usando la ecuación de Karplus-Altona⁴⁰. Los valores de la distribución se calcularon de acuerdo con la ecuación 2.1:

$$J = \sum P_{\phi, \psi} \times J_{i\phi, \psi} \quad \text{Ecuación 2.1}$$

Las distancias interprotónicas medias se calcularon aplicando la siguiente expresión:

$$\left\langle r^{-6} \right\rangle_{kl}^{-1/6} = \sum P_{\phi,\psi} \times r_{kl(\phi,\psi)}^{-6} \quad \text{Ecuación 2.2}$$

Las intensidades teóricas de los NOEs se determinaron de acuerdo con la matriz de relajación completa, empleando el programa NOEPROM⁴⁴, desarrollado en nuestro grupo de trabajo. Se asumió para ello un movimiento isotrópico de las moléculas y un valor de relajación externa de 0.1 s⁻¹. Se fijaron los valores de tiempo de correlación de 75 y 70 ps para **1** y **2** respectivamente, por ser los valores que mejor ajustaban los NOEs intra-residuos Gal1-Gal3 y Gal1-Gal5, que no dependen del valor de los ángulos glicosídicos.

5.4. Experimentos de RMN:

Los espectros del compuesto **1** se adquirieron en un instrumento Bruker de 500 MHz, equipado con una criosonda. La concentración del compuesto fue de 18 mM. Todos los experimentos se llevaron a cabo a 295 K en D₂O, utilizando como referencia externa DSS. Los espectros del compuesto **2** se adquirieron en un instrumento Bruker de 500 MHz, equipado con una sonda inversa ATM. La concentración del compuesto fue de 1 mM. Todos los experimentos se llevaron a cabo a 298 K en D₂O, utilizando como referencia externa DSS, y metanol-d₄, fijando la señal de residual del metanol a 3.31 ppm.

Para la asignación de los compuestos **1** y **2** se registraron experimentos de ¹H RMN, COSY⁴⁵, HSQC⁴⁶ y 1D-TOCSY⁴⁷ (este último con 80 ms de tiempo de mezcla). Para la determinación de la estructura tridimensional, se emplearon experimentos 1D y 2D-NOESY⁴⁸ y, en el caso del *O*-disacáridos, también se realizó un experimentos 2D-HSQC-NOESY⁴⁹. En todos los experimentos 2D, se seleccionó 500 ms como tiempo de mezcla, mientras que para los experimentos 1D se seleccionaron tiempos de mezcla de 500 y 300 ms.

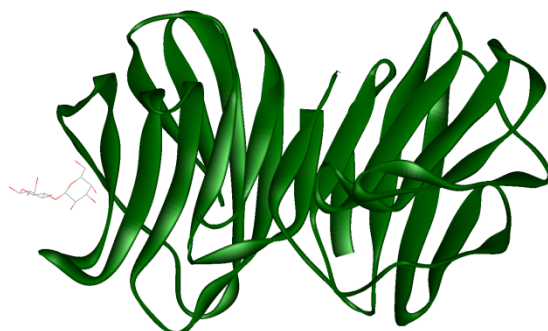
6. Bibliografía:

- (1) Dam, T. K.; Brewer, C. F. *Chem. Rev.* **2002**, *102*, 387.
- (2) Gabius, H. J. *Pharmacol. Res.* **1998**, *15*, 23.
- (3) Mikkelsen, L. M.; Hernaiz, M. J.; Martín-Pastor, M.; Skrydstrup, T.; Jimenez-Barbero, J. *J. Am. Chem. Soc.* **2002**, *124*, 14940.
- (4) Weatherman, R. V.; Kiessling, L. L. *J. Org. Chem.* **1996**, *61*, 534.
- (5) Weatherman, R. V.; Mortell, K. H.; Chervenak, M.; Kiessling, L. L.; Toone, E. *J. Biochemistry (Mosc.)* **1996**, *35*, 3619.
- (6) Searle, M. S.; Williams, D. H. *J. Am. Chem. Soc.* **1992**, *114*, 10690.
- (7) Wu, T. C.; Goekjian, P. G.; Kishi, Y. *J. Org. Chem.* **1987**, *52*, 4819.
- (8) Babirad, S. A.; Wang, Y.; Goekjian, P. G.; Kishi, Y. *J. Org. Chem.* **1987**, *52*, 4825.
- (9) Goekjian, P. G.; Wu, T. C.; Kang, H. Y.; Kishi, Y. *J. Org. Chem.* **1987**, *52*, 4823.
- (10) Goekjian, P. G.; Wu, T. C.; Kang, H. Y.; Kishi, Y. *J. Org. Chem.* **1991**, *56*, 6422.
- (11) Goekjian, P. G.; Wu, T. C.; Kishi, Y. *J. Org. Chem.* **1991**, *56*, 6412.
- (12) Haneda, T.; Goekjian, P. G.; Kim, S. H.; Kishi, Y. *J. Org. Chem.* **1992**, *57*, 490.
- (13) Miller, W. H.; Ryckman, D. M.; Goekjian, P. G.; Wang, Y.; Kishi, Y. *J. Org. Chem.* **1988**, *53*, 5580.
- (14) O'Leary, D. J.; Kishi, Y. *J. Org. Chem.* **1993**, *58*, 304.
- (15) O'Leary, D. J.; Kishi, Y. *J. Org. Chem.* **1994**, *59*, 6629.
- (16) Wang, Y.; Goekjian, P. G.; Ryckman, D. M.; Kishi, Y. *J. Org. Chem.* **1988**, *53*, 4151.
- (17) Wang, Y.; Goekjian, P. G.; Ryckman, D. M.; Miller, W. H.; Babirad, S. A.; Kishi, Y. *J. Org. Chem.* **1992**, *57*, 482.
- (18) Wei, A.; Haudrechy, A.; Audin, C.; Jun, H.-S.; Haudrechy-Bretel, N.; Kishi, Y. *J. Org. Chem.* **1995**, *60*, 2160.
- (19) Wei, A.; Kishi, Y. *J. Org. Chem.* **1994**, *59*, 88.
- (20) Wang, Y.; Babirad, S. A.; Kishi, Y. *J. Org. Chem.* **1992**, *57*, 468.
- (21) Skrydstrup, T.; Mazéas, D.; Elmouchir, M.; Doisneau, G.; Riche, C.; Chiaroni, A.; Beau, J.-M. *Chem.--Eur. J.* **1997**, *3*, 1342.
- (22) Neuman, A.; Longchambon, F.; Abbes, O.; Gillier-Pandraud, H.; Perez, S.; Rouzaud, D.; Sinay, P. *Carbohydr. Res.* **1990**, *195*, 187.
- (23) Bock, K.; Duus, J. Ø.; Refn, S. *Carbohydr. Res.* **1994**, *253*, 51.
- (24) Geyer, A.; Hummel, G.; Eisele, T.; Reinhardt, S.; Schmidt, R. R. *Chem.--Eur. J.* **1996**, *2*, 981.
- (25) Espinosa, J.-F.; Dietrich, H.; Martín-Lomas, M.; Schmidt, R. R.; Jiménez-Barbero, J. *Tetrahedron Lett.* **1996**, *37*, 1467.
- (26) Espinosa, J.-F.; Martín-Pastor, M.; Asensio, J. L.; Dietrich, H.; Martín-Lomas, M.; Schmidt, R. R.; Jiménez-Barbero, J. *Tetrahedron Lett.* **1995**, *36*, 6329.
- (27) Rubinstenn, G.; Sinay, P.; Berthault, P. *J. Phys. Chem. A* **1997**, *101*, 2536.
- (28) Espinosa, J.-F.; Bruix, M.; Jarreton, O.; Skrydstrup, T.; Beau, J.-M.; Jiménez-Barbero, J. *Chem.--Eur. J.* **1999**, *5*, 442.
- (29) Asensio, J. L.; Cañada, F. J.; Cheng, X.; Khan, N.; Mootoo, D. R.; Jiménez-Barbero, J. *Chem.--Eur. J.* **2000**, *6*, 1035.

- (30) Jimenez-Barbero, J.; Espinosa, J.-F.; Asensio, J. L.; Cañada, F. J.; Poveda, A. *Advances in Carbohydrate Chemistry and Biochemistry*; Horton, D. ed.; Academic Press, 2001; Vol. 56.
- (31) Asensio, J. L.; Martín-Pastor, M.; Jiménez-Barbero, J. *J. Mol. Struct. Theochem* **1997**, 395-396, 245.
- (32) Dabrowski, J.; Kozar, T.; Grosskurth, H.; Nifant'ev, N. E. *J. Am. Chem. Soc.* **1995**, 117, 5534.
- (33) Dowd, M. K.; French, A. D.; Reilly, P. J. *Carbohydr. Res.* **1992**, 233, 15.
- (34) Imberty, A.; Mikros, E.; Koca, J.; Mollicone, R.; Oriol, R.; Perez, S. *Glycoconj. J.* **1995**, 12, 331.
- (35) Kony, D.; Damm, W.; Stoll, S.; Hunenberger, P. H. *J. Phys. Chem. B* **2004**, 108, 5815.
- (36) Cheetham, N. W.; Dasgupta, P.; Ball, G. E. *Carbohydr. Res.* **2003**, 338, 955.
- (37) French, A. D.; Kelterer, A. M.; Johnson, G. P.; Dowd, M. K.; Cramer, C. J. *J. Mol. Graph. Model.* **2000**, 18, 95.
- (38) Boelens, R.; Koning, T. M. G.; Kaptein, R. *J. Mol. Struct.* **1988**, 173, 299.
- (39) Boelens, R.; Koning, T. M. G.; Vandermarel, G. A.; Vanboom, J. H.; Kaptein, R. *J. Magn. Reson.* **1989**, 82, 290.
- (40) Altona, C.; Sundaral, M. *J. Am. Chem. Soc.* **1972**, 94, 8205.
- (41) Espinosa, J. F.; Canada, F. J.; Asensio, J. L.; Martín-Pastor, M.; Dietrich, H.; Martín-Lomas, M.; Schmidt, R. R.; Jimenez-Barbero, J. *J. Am. Chem. Soc.* **1996**, 118, 10862.
- (42) Lopez, R.; Fernandez-Mayoralas, A. *J. Org. Chem.* **1994**, 59, 737.
- (43) Vauzeilles, B.; Sinaÿ, P. *Tetrahedron Lett.* **2001**, 42, 7269.
- (44) Martín-Pastor, M., <http://desoft03.usc.es/mmartin/software.html>
- (45) Aue, W. P.; Bartholdi, E.; Ernst, R. R. *J. Chem. Phys.* **1976**, 64, 2229.
- (46) Willker, W.; Leibfritz, D.; Kerssebaum, R.; Bermel, W. *Magn. Reson. Chem.* **1993**, 31, 287.
- (47) Adell, P.; Parella, T.; Sanchezferrando, F.; Virgili, A. *J. Magn. Reson. Series B* **1995**, 108, 77.
- (48) Neuhaus, D. W., M. P. *The Nuclear Overhauser Effect in Structural and Conformational Analysis*; VCH: New York, 1989.
- (49) Ronald Wagner, S. B. *Magn. Reson. Chem.* **1997**, 35, 199.

Capítulo 3

Estudio del β -C-Gal(1 \rightarrow 3)- β -Glc-OMe en presencia de Galectina-1: comparativa entre el *O*- y el *C*-disacárido.



1. Introducción y antecedentes:

El papel crucial de las interacciones proteína-carbohidrato en numerosos mecanismos biológicos, incluyendo la adhesión celular, crecimiento celular, inflamación y respuesta inmune, ha provocado un creciente interés en el análisis estructural de los complejos lectina-carbohidrato¹. Las lectinas, como se ha mencionado en la introducción, son proteínas que se unen específicamente a los carbohidratos participando en procesos de reconocimiento biológico. Dentro de las lectinas se encuentra un grupo de proteínas llamadas galectinas², denominadas originalmente como lectinas de tipo-S para denotar la presencia de puentes disulfuro y residuos de cisteína, y que se caracterizan por presentar interacciones específicas con β -galactósidos. La primera galectina encontrada en mamíferos fue la galectina-1, aislada en 1976 del músculo de pollo y del corazón y pulmón de ternero³⁻⁴. Su tamaño es de aproximadamente 15 kD, encontrándose como un dímero no covalente. A principios de los años 80, otra proteína de 35 kD, conocida ahora como galectina-3, que también interactuaba con β -galactósidos, fue hallada en los fibroblastos de un ratón⁵. Hasta la fecha, se han identificado hasta 15 tipos de galectinas en mamíferos⁶.

La galectina-1 ha mostrado jugar un papel importante en la apoptosis de los linfocitos T⁷, activada al unirse a los azúcares presentes en la superficie de estas células del sistema inmune. La galectina-1 se expresa en diversos tipos de tumores, tales como carcinoma de mama, próstata, ovario y melanoma. De hecho, se ha demostrado la relación entre los niveles de expresión de esta proteína y la agresividad de dichos tumores⁸. A partir de estas observaciones, se postuló la hipótesis de que la expresión de galectina-1 podría serle útil a la célula tumoral como mecanismo para eliminar linfocitos T, eludiendo así las defensas de nuestro organismo. Por tanto, la inhibición de la galectina-1 podría constituir una posible vía terapéutica para evitar la extensión de diversos tipos de cáncer⁹⁻¹⁰.

La interacción entre la galectina-1 y diversos β -galactósidos ha sido estudiada por diversas técnicas, entre las que se encuentran la cristalografía de rayos X¹¹⁻¹⁵, la calorimetría isoterma de titulación¹⁶⁻¹⁸, los cálculos de dinámica molecular^{17,19-21}, la

espectroscopía UV-Raman²² y la resonancia magnética nuclear^{17,20-21,23-25}. En la **Tabla 3.1** se recoge un resumen de los carbohidratos estudiados, indicándose la(s) técnica(s) utilizadas en cada caso.

Ligando	Tipo de unión	Técnicas			
		Rayos-X	ITC*	NMR	MD
β -D-Galactosa (1)		1W6M ¹³		✓ ¹⁷	✓ ¹⁷
Phenyl- β -D-Galactósido (2)				✓ ²⁵	
Galacto- <i>N</i> -biósido (3)				✓ ¹⁷	✓ ¹⁷
β -Lactosa (4)	$\beta(1-4)\text{Glc}$	1GZW ¹³ 1W6O ¹³ 2ZKN ¹⁵	✓ ^{16,18}	✓ ^{17,20,23}	✓ ^{17,19-20}
β -Lactosa-OMe (5)	$\beta(1-4)\text{Glc}$		✓ ¹⁸	✓ ²⁵	
<i>N</i> -Acetil-Lactosamina (LacNac-II) (6)	$\beta(1-4)\text{Glc}$	1GAN ¹⁴ 1SLT ¹² 1W6P ¹³	✓ ^{16,18}	✓ ^{17,25}	✓ ^{17,19}
Gal $\beta(1-4)$ Fruc (7)	$\beta(1-4)\text{Fruc}$		✓ ¹⁸		
Gal $\beta(1-4)$ Man (8)	$\beta(1-4)\text{Man}$		✓ ¹⁸		
MeO-2-Gal $\beta(1-4)$ Glc (9)	$\beta(1-4)\text{Glc}$		✓ ¹⁸		
Glycoconjugados <i>N</i> -Acetil-lactosamina	$\beta(1-4)\text{Glc}$	1SLA ¹¹ 1SLB ¹¹ 1SLC ¹¹			
DiLacNac	$\beta(1-4) \beta(1-3) \beta(1-4)$		✓ ¹⁶		✓ ¹⁹
TriLacNac	$\beta(1-4) \beta(1-3) \beta(1-4)$ $\beta(1-3) \beta(1-4)$		✓ ¹⁶		
3'OSO3-LacNacII (10)	$\beta(1-4)\text{Glc}$				✓ ¹⁹
Neu5Ac- $\alpha(2-3)$ Gal $\beta(1-4)$ GlcNac	$\beta(1-4)\text{Glc}$		✓ ¹⁶		✓ ¹⁹
Neu5Ac $\alpha(2-3)$ Gal $\beta(1-4)$ Glc	$\beta(1-4)\text{Glc}$				✓ ²¹
β -Gal $\beta(1-4)$ Xyl-OMe (11)	$\beta(1-4)\text{Xyl}$			✓ ²³	
Lacto- <i>N</i> -Biósido (LacNac-I) (12)	$\beta(1-3)\text{Glc}$		✓ ^{16,18}	✓ ^{17,25}	✓ ^{17,19}
Lacto- <i>N</i> -Tetraosa	$\beta(1-3) \beta(1-3) \beta(1-4)$		✓ ¹⁶		
Gal $\beta 1-3$ GalNac (13)	$\beta(1-3)\text{Gal}$				✓ ²¹
Gal $\beta(1-3)$ Ara (14)	$\beta(1-3)\text{Ara}$		✓ ¹⁸		
β -Gal $\beta(1-3)$ Xyl-OMe (15)	$\beta(1-3)\text{Xyl}$			✓ ²³	
Tiodigalactósido (16)	$\beta(1-1)\text{Glc}$	1A78 ¹⁴	✓ ^{16,18}		
C-Lactosa (β -C-Gal $\beta(1-4)$ Glc-OMe) (17)	$\beta(1-4)\text{Glc}$			✓ ^{23,26}	
C-Lactosamina (β -C-Gal $\beta(1-4)$ GlcNac-OMe) (18)	$\beta(1-4)\text{Glc}$			✓ ²⁷	

Tabla 3.1: Carbohidratos estudiados frente a la galectina-1, mediante distintas técnicas.

Las estructuras de Rayos X de la galectina-1 natural y mutada complejada con galactosa¹³, lactosa (Lac)^{13,15}, tiodigal (*S*- $\beta\text{Gal}\beta(1\rightarrow1)\text{Gal}$)¹⁴, *N*-acetil-lactosamina (NacLac)¹²⁻¹⁴, o con glicoconjugados de NacLac¹¹, han revelado que la proteína es

un dímero, formado por dos monómeros idénticos relacionados por un eje de simetría C₂. Cada unidad esta compuesta por dos láminas beta antiparalelas de 5 o 6 hebras, que se disponen formando un sandwich beta. El centro de reconocimiento del carbohidrato se encuentra situado en la parte opuesta del dímero, entre la quinta y sexta hebra, y la interacción con el carbohidrato se produce a través de enlaces de hidrógeno y de *stacking* con un residuo de triptófano altamente conservado. En general, la estructura abierta del centro de reconocimiento permite la unión de glicoconjugados con unidades de galactosa terminales. En todos los disacáridos analizados, la conformacion bioactiva de los enlaces $\beta(1\rightarrow4)$ corresponde a la familia conformacional *syn*, que es la conformación mayoritaria en disolución.

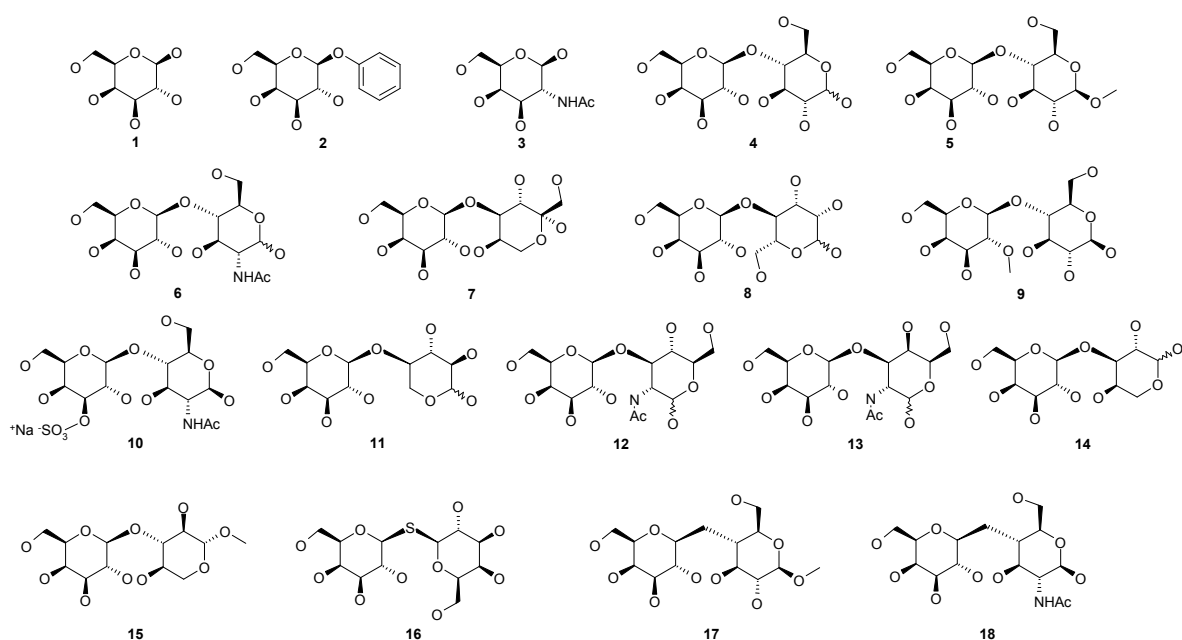


Figura 3.1: Estructura de algunos de los carbohidratos incluidos en la tabla anterior.

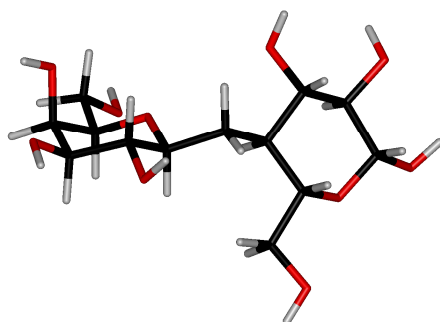


Figura 3.2: Conformación *syn* en el caso de la C-lactosa.

Por otra parte, estudios de calorimetría isoterma de titulación (ITC) han demostrado que la galactosa se une a la galectina-1 con baja afinidad, por lo que las interacciones que se establecen entre la unidad reductora del disacárido y la proteína son necesarias para aumentar la afinidad del ligando¹⁷. Además, estos estudios pusieron de manifiesto que la galectina-1 no sólo reconoce β -galactósidos con enlaces glicosídicos $\beta(1\rightarrow4)$, sino también aquellos con enlaces $\beta(1\rightarrow3)$ como lacto-*N*-biosa o $\text{Gal}\beta(1\rightarrow3)\text{Ara}$ ¹⁸. Como en otros complejos proteína-carbohidrato, el proceso de reconocimiento está favorecido entálpicamente y presenta el conocido efecto de la compensación entalpía-entropía.

Como se describe en la introducción de esta memoria, la RMN es una técnica muy útil para el estudio del reconocimiento carbohidrato-proteína gracias a su gran versatilidad, siendo posible desde determinar la estructura tridimensional del complejo hasta simplemente confirmar la interacción entre la proteína y el carbohidrato, y de hecho la interacción entre la galectina-1 y la lactosa también ha sido estudiada mediante RMN²⁶. Así, estudios de diferencia de transferencia de saturación (STD) han demostrado que la mayor transferencia de magnetización entre la galectina-1 y varios β -galactósidos ocurre en los protones en posiciones 4 y 6 del anillo de galactosa y, en menor medida, en el protón en posición 5, por lo que estos protones son los que establecen los contactos más íntimos con la proteína²⁵. Por otro lado, la asignación completa de las resonancias de ^1H , ^{13}C y ^{15}N de galectina-1 doblemente marcada²⁴ ha permitido el estudio detallado del complejo con lactosa en disolución²⁰. Este análisis ha revelado que la lactosa se une a los dos centros de reconocimiento con cooperatividad negativa, es decir que la primera unidad de lactosa se une más fuertemente a la galectina-1 que la segunda.

El NOE transferido (tr-NOE) es otro tipo de experimentos de RMN que permite determinar la conformación de un ligando unido a una proteína²⁸. Esta técnica se basa en la dependencia de la intensidad de NOE con el tiempo de correlación, que a su vez es función del tamaño molecular. Cuando un ligando de bajo peso molecular está libre en disolución, los NOEs entre sus protones son positivos y de intensidad débil, mientras que cuando se une a la proteína, el ligando se comporta como una macromolécula que se caracteriza por NOEs negativos e intensos. Si la velocidad de intercambio entre la forma libre y complejada es suficientemente rápida en la escala de tiempo de la relajación, situación frecuente en

complejos proteína-carbohidrato^{1,29-30}, los NOEs observados entre los protones del ligando son un promedio entre los estados libre y asociado, pero que viene principalmente determinado por el estado asociado, por lo que proporcionan información de la conformación bioactiva.

La técnica de tr-NOE ha sido previamente utilizada por nuestro grupo de investigación para determinar la conformación de lactosa y Gal β (1 \rightarrow 4)Xyl en el centro de reconocimiento de la galectina-1²⁷. En los espectros tr-NOESY y tr-ROESY de estos dos disacáridos en presencia de galectina-1 sólo se detectaron los NOEs característicos de la conformación *syn*, indicando que esta familia conformacional es reconocida por la proteína de acuerdo con lo encontrado para otros galactosidos β (1 \rightarrow 4) mediante cristalografía de rayos-X. Como se ha comentado en capítulos previos, la conformación *syn* es muy mayoritaria en disacáridos β (1 \rightarrow 4), por lo que el reconocimiento exclusivo de esta conformación podría deberse a la alta energía del resto de conformaciones, que apenas se encuentran pobladas en disolución. Los estudios de tr-NOE de los derivados C-glicosídicos de lactosa y lactosamina, que presentan una mayor flexibilidad en torno a la unión glicosídica con tres conformaciones diferentes pobladas en disolución (*syn*, *anti- ψ* y *anti- ϕ*)^{23,26-27}, demostraron que estos dos C-disacáridos son reconocidos por la galectina-1 en su conformación *syn* mientras que la unión del mínimo global, *anti- ψ* , a la proteína no fue detectada. Esto pone de manifiesto que las características topológicas del centro de reconocimiento de la proteína favorecen la unión de la conformación *syn*, pero no de la *anti* y que, por tanto, el reconocimiento exclusivo de este conformero por la galectina-1 constituye un ejemplo de selección conformacional.

Como se ha mencionado anteriormente, distintos estudios de calorimetría han demostrado que los galactósidos β (1 \rightarrow 3) son también ligandos de la galectina-1^{16,18}. Sin embargo, hasta la fecha, no existen estudios de rayos-X de este tipo de complejos. El grupo de Bourne propuso un modo de unión a la galectina-1 basado en la estructura de rayos-X del complejo con *N*-acetyl lactosamina¹¹. En este disacárido, el hidroxilo en posición 3 del anillo de GlcNAc tiene un papel muy importante, ya que forma un enlace de hidrógeno con un aminoácido de la galectina, interacción que no es posible en el caso de la lacto-*N*-biosa con enlace β (1 \rightarrow 3) debido a la presencia del enlace glicosídico en esa posición. La propuesta de Bourne¹¹ sitúa el conformero

syn de mínima energía de la lacto-*N*-biosa con la unidad de galactosa ocupando la misma posición que en el caso de disacáridos $\beta(1\rightarrow4)$, pero con la unidad de glucosa girada 180° de tal modo que el hidroxilo O4 se situaría en una posición equivalente a la del hidroxilo O3 y estaría implicado en las mismas interacciones, lo que explicaría que ambos disacáridos tengan afinidades parecidas. Esta hipótesis está de acuerdo con simulaciones de dinámica molecular que predicen que la conformación *syn* de la lacto-*N*-biosa establece una red de enlaces de hidrógeno con los aminoácidos de la proteína similar a la de la *N*-acetil-lactosamina^{17,19}. Estudios de NOE transferido de la lacto-*N*-biosa, unidad constituyente del pentasacárido del gangliósido GM₁, en presencia de galectina-1 han confirmado que el mínimo *syn* es la conformación reconocida por la proteína²¹.

Dada la ausencia de estructuras de rayos-X de galectina-1 complejada con disacáridos $\beta(1\rightarrow3)$, el objetivo de este capítulo es determinar la conformación del disacárido Gal $\beta(1\rightarrow3)$ Glc β OMe y de su *C*-análogo que son reconocidas por la galectina-1, mediante experimentos de NOE transferido. Como se ha visto en el capítulo anterior, la familia conformacional *syn* es muy predominante en el *O*-disacárido, mientras que en el *C*-disacárido se produce una disminución de la población de este conformero y las dos familias conformacionales, *syn* y *anti-ψ*, se encuentran significativamente pobladas en disolución. Por tanto, el conocimiento de la conformación bioactiva del derivado *C*-glicosídico debería permitir determinar si los disacáridos con enlaces $\beta(1\rightarrow3)$ también son reconocidos mediante un proceso de selección conformacional, por parte de la lectina, como se ha visto para uniones $\beta(1\rightarrow4)$. Estos estudios de RMN se han complementado con un estudio computacional de *docking* utilizando la estructura de rayos-X del complejo galectina-lactosa como punto de partida, con el fin de caracterizar las interacciones proteína-carbohidrato que estabilizan el complejo.

2. Análisis conformacional del β -Gal(1 \rightarrow 3) β -Glc-OMe (1) en presencia de galectina-1:

En la **Figura 3.3** se pueden observar los espectros de protón del β -Gal(1 \rightarrow 3) β -Glc-OMe (1) en ausencia y presencia de galectina-1. Tras la adición de la proteína, se observó un ensanchamiento de las señales, sobre todo de aquellas correspondientes al anillo de galactosa. Este fenómeno es consecuencia del acortamiento de los tiempos de relajación T_2 , que se produce debido al equilibrio entre los estados libre del disacárido y asociado a la lectina, y que permite confirmar que el disacárido interacciona con ésta. El hecho de que este efecto sea más pronunciado en el anillo de galactosa refleja que esta unidad presenta menos movilidad en el complejo que la unidad de glucosa. Como solamente se observa un único conjunto de señales, se concluye que el intercambio entre la forma libre y complejada es rápido en la escala de tiempo de los desplazamientos químicos.

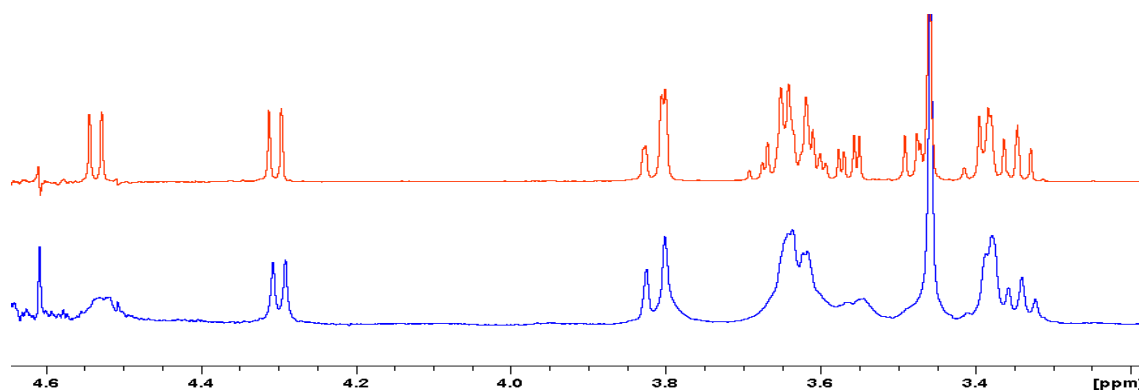


Figura 3.3: Espectro de protón de **1** en presencia (azul) y ausencia (rojo) de galectina-1

Tras este resultado inicial, y para poder determinar la conformación bioactiva del ligando, se realizaron experimentos de NOE transferido. Mientras que en el espectro NOESY del disacárido libre, los picos de cruce entre los protones del ligando fueron de signo contrario a la diagonal (positivos), los picos de cruce en el espectro tr-NOESY fueron del mismo signo (negativos). Este cambio de signo constituye la segunda evidencia experimental de la interacción entre la proteína y el ligando.

Debido al solapamiento del espectro, se realizó un tr-HSQC-NOESY para conseguir mayor dispersión de señales, tal como se había hecho anteriormente para el

disacárido libre en agua. Los espectros HSQC-NOESY y tr-HSQC-NOESY son muy parecidos (**Figura 3.4**) salvo por la menor relación señal:ruido del espectro en presencia de proteína, lo que provoca la desaparición de algunos de los NOEs débiles que se observaban en el espectro del ligando libre. La similitud entre los espectros sugiere que el mínimo *syn*, que es muy predominante en disolución, es la conformación reconocida por la galectina-1.

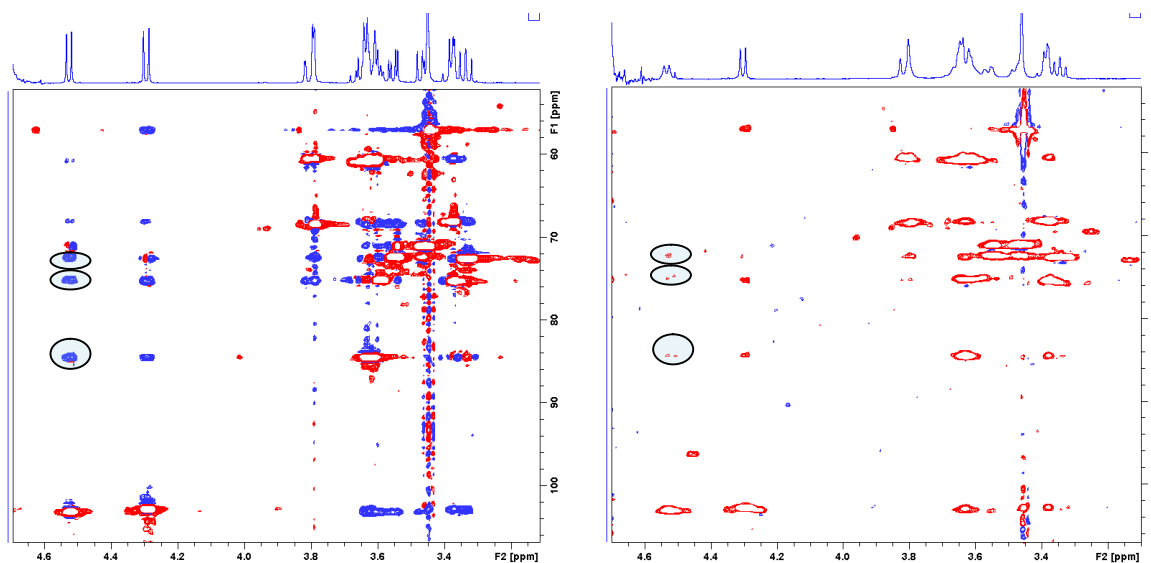


Figura 3.4: Espectro 2D-HSQC-NOESY (izquierda) y tr-2D-HSQC-NOESY (derecha)

En el capítulo 2 se vio que los NOEs más informativos del *O*-disacárido son los que se dan entre el H1 del residuo de galactosa y los protones de la glucosa. El único NOE interglucosídico que se observa en el experimento tr-HSQC-NOESY es el NOE Gal1-Glc3, exclusivo del conformero *syn*, que presenta una intensidad comparable a los NOE intra-glicosídicos Gal1-Gal3 y Gal1-Gal5, lo que confirma que la galectina-1 reconoce al conformero *syn* de **1**. El hecho de que el NOE Gal1-Glc4, exclusivo del conformero *anti-ψ*, no se detecte en el espectro tr-HSQC-NOESY sugiere que esta conformación no se une a la galectina-1, ya que los protones H1-H4 estarían a una distancia muy corta en la conformación bioactiva y el NOE correspondiente debería ser de intensidad parecida a los NOEs intra-glicosídicos, aunque no podemos descartar que la ausencia de este NOE sea debida a la falta de sensibilidad del experimento.

3. Análisis conformacional del β -C-Gal(1 \rightarrow 3) β -Glc-OMe (**2**) en presencia de galectina-1:

De igual manera que en **1**, la adición de galectina-1 a la disolución de β -C-Gal(1 \rightarrow 3) β -Glc-OMe (**2**), produjo un ensanchamiento de las señales en el espectro de ^1H , sobre todo en aquellas correspondientes al anillo de galactosa (**Figura 3.5**) y un cambio de signo en los NOEs en el experimento NOESY, lo que indicó que el C-disacárido es también reconocido por la galectina-1. La adición de un exceso de O-disacárido produce el afinamiento de las señales del C-disacárido, lo que indica que ambos compuestos compiten por el mismo centro de reconocimiento y que la afinidad del C-disacárido por la proteína es menor.

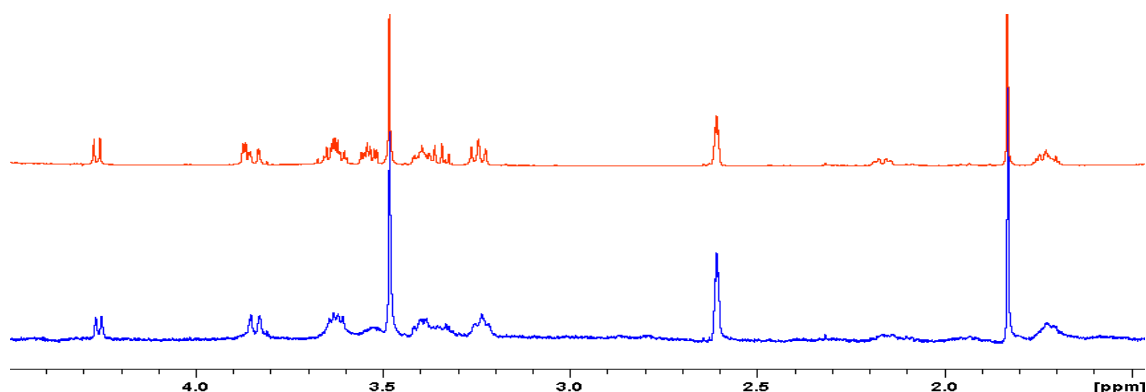


Figura 3.5: Espectro de protón de **2** en presencia (azul) y ausencia (rojo) de galectina-1

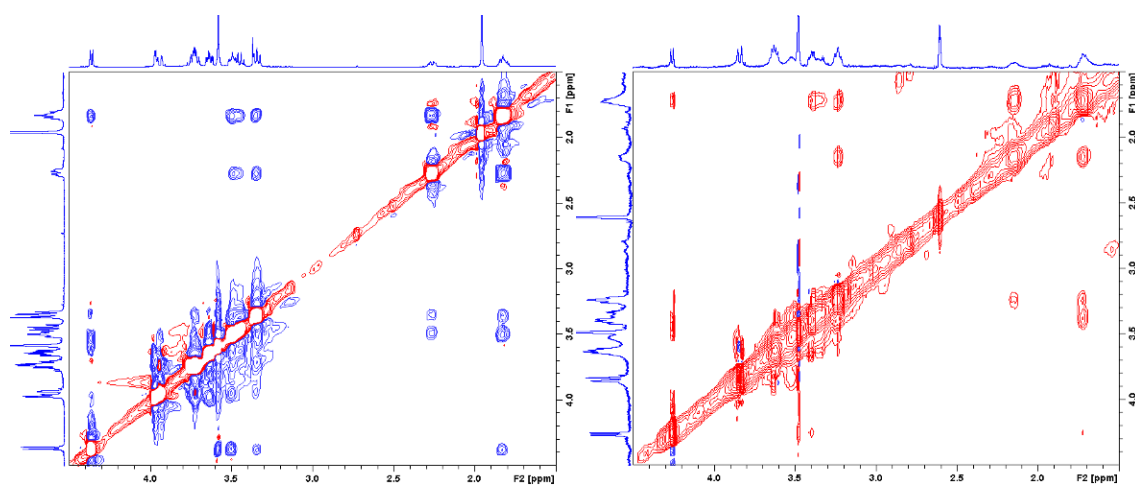


Figura 3.6: Espectro 2D-NOESY (izquierda) y tr-2D-NOESY (derecha)

En este caso, no fue posible confirmar la presencia inequívoca de los NOEs exclusivos Gal1-Glc3 y Gal1-Glc4 debido al alto solapamiento de las señales. Asimismo, la poca cantidad de muestra disponible impidió utilizar el experimento HSQC-NOESY, por lo que el análisis se centró en los NOEs entre los protones metilénicos del puente carbonado y los protones de la glucosa, ya que la intensidad de estos NOEs también puede relacionarse con las poblaciones de los conformeros *syn* y *anti*- ψ y aparecen en una zona del espectro con menor solapamiento (**Figura 3.7**).

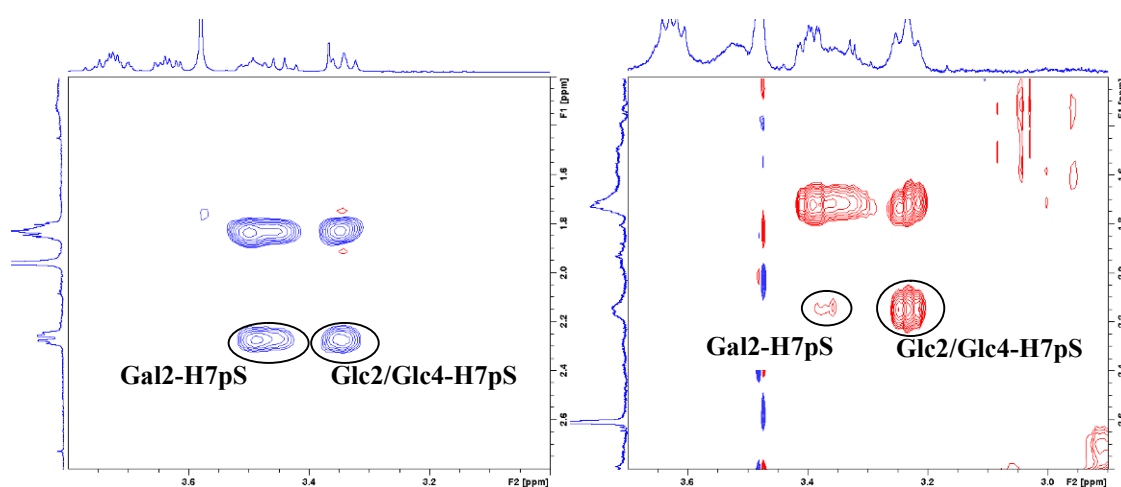


Figura 3.7: Expansión de los espectros 2D-NOESY (izquierda) y tr-2D-NOESY (derecha), mostrando los NOEs de interés.

Así, si se comparan los espectros NOESY y tr-NOESY, puede observarse un aumento claro de la intensidad del NOES H7pS-Glc2/Glc4 con respecto al NOE H7pS-Gal2, indicando un cambio conformacional debido a la interacción con la proteína. Así, mientras que la distancia H7pS y Gal2 es similar en el conformero *syn*- ψ (+) y *anti*- ψ (en torno a 3 Å), ya que depende el ángulo diedro ϕ , la distancia entre H7pS y los protones Glc2 y Glc4 es mucho menor en el conformero *syn*- ψ (+) (en torno a 2.5 Å en ambos casos) que en el conformero *anti*- ψ (3.9 y 3.2 Å respectivamente). Por tanto, el cambio en la intensidad relativa de los NOEs sugiere que el equilibrio conformacional se desplaza hacia la conformación *syn*- ψ (+) en el estado asociado.

Además, se realizaron experimentos tr-ROESY para excluir efectos de difusión de espín mediados por la proteína (en este experimento los NOEs aparecen de signo contrario a la diagonal –positivos–, mientras que los efectos indirectos son débiles y

de signo contrario). Como puede apreciarse en la **Figura 3.8**, los NOEs utilizados para el análisis son positivos y sus intensidades relativas similares a las observadas en el tr-NOESY, confirmando que se trata de efectos directos.

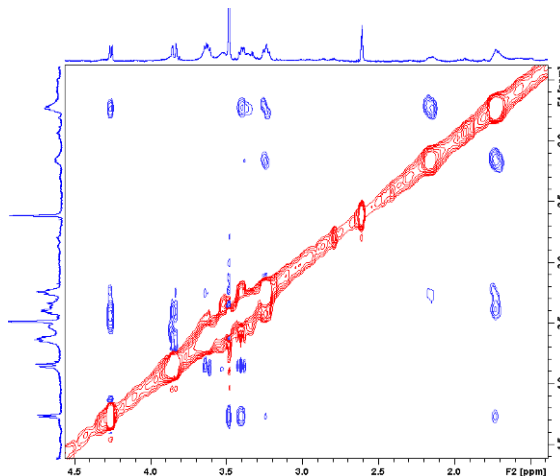


Figura 3.8: Espectro tr-2D-ROESY.

4. Estudio de acoplamiento o docking:

El estudio computacional de *docking* fue realizado por el grupo de Sonsoles Martín Santamaría utilizando el programa AutoDock 3.1³¹, siguiendo el protocolo descrito en la parte experimental. Con el objeto de validar el procedimiento computacional, se llevó a cabo el estudio de *docking* con la lactosa, utilizando la estructura de rayos-X del complejo glectina-1-lactosa (PDB 1GZW) como punto de partida¹³. Para la lactosa, el programa AutoDock fue capaz de reproducir la conformación de unión cristalográfica con una energía de unión de $-3.88 \text{ kcal mol}^{-1}$, obteniéndose un rango de energías para la totalidad de las soluciones que osciló entre -4.51 y $-2.62 \text{ kcal mol}^{-1}$. Las interacciones más importantes fueron el enlace de hidrógeno entre el O4 de la unidad de galactosa y los aminoácidos His-44 y Arg 48 y el apilamiento (“*stacking*”) entre los C3, C4 y C5 de la galactosa y el anillo aromático del Trp-68 (**Figura 3.9**). La interacción más destacada con el anillo de glucosa fue el enlace de hidrógeno entre O3 y los aminoácidos Arg-48 y Glu-71.

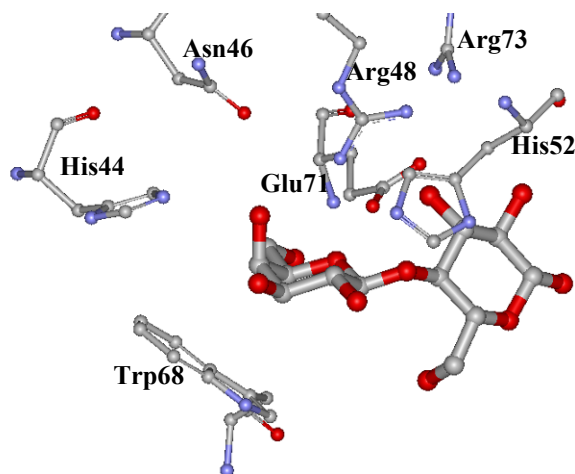


Figura 3.9: Conformación *syn-ψ(+)* de la lactosa en la estructura cristalográfica 1GZW.

Una vez que se comprobó que el programa AutoDock reproducía los datos experimentales de rayos-X, se realizó el *docking* del compuesto **1**. El método computacional predijo un valor de energía de unión de $-3.58 \text{ kcal mol}^{-1}$ para el complejo en el que el disacárido adopta una conformación *syn-ψ(+)*. Este conformero se acopla perfectamente en el sitio de reconocimiento de la glectina-1,

Si se compara la disposición de este compuesto **1** con respecto a la estructura cristalográfica del complejo galectina-lactosa, puede observarse (**Figura 3.11**) que la desviación cuadrática media de la posición del anillo de galactosa es de 0.7 Å, mientras que las interacciones clave se mantienen básicamente idénticas

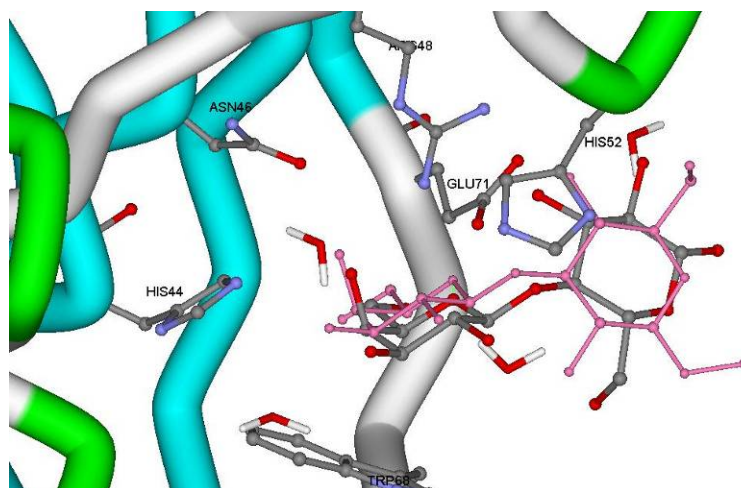


Figura 3.11: Superposición de la conformación *syn-ψ(+)* del compuesto **1** (en rosa) con la estructura cristalográfica de la lactosa (código CPK) en el sitio de reconocimiento de la galectina-1.

Respecto al compuesto *C*-glicosídico, el análisis computacional de **2** reveló la existencia de tres familias conformacionales en disolución con una población mayor del 3%. La conformación mayoritaria se corresponde con la disposición *syn-ψ(+)*, mientras que las conformaciones *anti-ψ* y *syn-ψ(-)* son minoritarias. Estas tres conformaciones fueron empleadas como geometrías de partida para llevar a cabo el estudio de acoplamiento o *docking* en el sitio de reconocimiento de la galectina-1.

Las tres conformaciones de partida proporcionaron soluciones razonables tras el proceso de acoplamiento a la galectina-1. De acuerdo con las energías de interacción proporcionadas por el programa AutoDock, las conformaciones de unión *syn-ψ(+)* (energía de unión de -3.67 kcal mol⁻¹) y *syn-ψ(-)* (-3.69 kcal mol⁻¹) son las más favorables, seguidas por la orientación *anti-ψ* (-2.98 kcal mol⁻¹). La disposición obtenida para el compuesto **2** en su conformación *syn-ψ(+)* es muy similar a la descrita anteriormente para el análogo *O*-glicosídico. Como se puede ver en las **Figuras 3.12 y 3.13**, la superposición de ambas estructuras es casi perfecta.

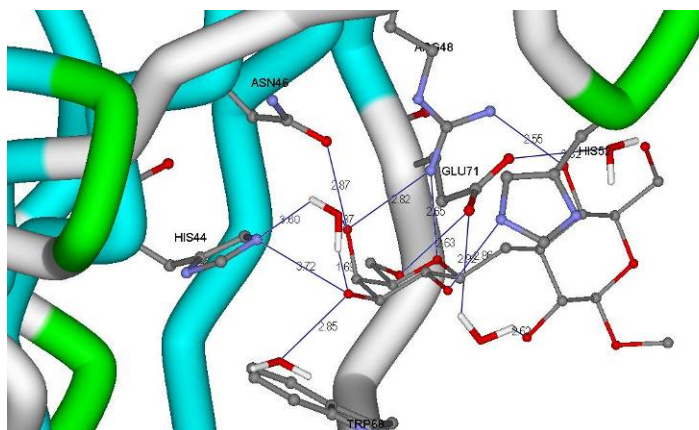


Figura 3.12: Distancias claves entre el ligando en conformación *syn-ψ(+)* y los aminoácidos del sitio de reconocimiento

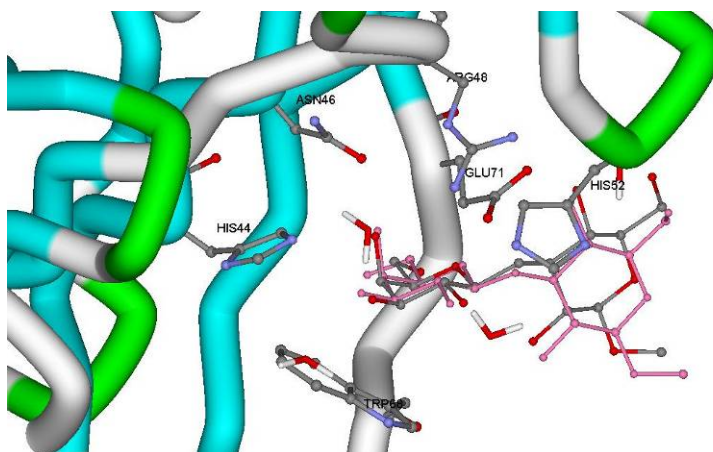


Figura 3.13: Superposición del compuesto **1** (en rosa) y *C*-glicósido (**2**) (en código CPK), ambas en conformación *syn-ψ(+)*.

Las disposiciones obtenidas para el compuesto *C*-glicósido en sus otras conformaciones son similares a la lactosa respecto a la posición y orientación del residuo de galactosa, mientras que las interacciones del residuo de Glc con los aminoácidos del centro de reconocimiento son diferentes para cada cónfórmero. En la **Figura 3.14** puede verse la superposición de los cónfórmeros *syn-ψ(+)* y *syn-ψ(-)*. En este caso, las interacciones con el residuo de glucosa son poco significativas, ya que el enlace de hidrógeno entre O4 de Glc y la Arg48 o entre O6 de Glc y la His52 no pueden tener lugar. En este caso, sólo son posibles interacciones de enlace de hidrógeno mediadas por una molécula de agua entre el O4 y el O6 de Glc y el Glu71. Aunque la orientación de este enlace no es del todo óptima, también podría darse un enlace de hidrógeno adicional entre el O2 de Glc y la His52.

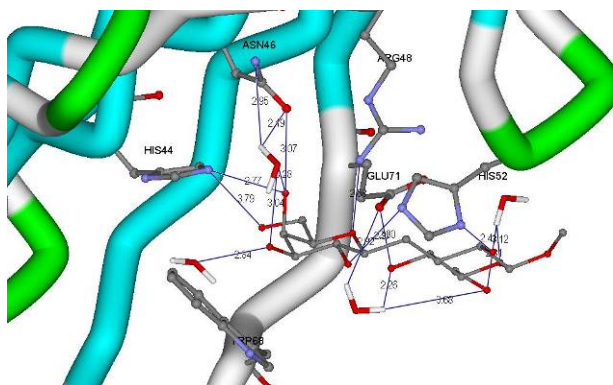


Figura 3.14: Distancias clave entre el ligando en conformación *syn-ψ(-)* y los aminoácidos del sitio de reconocimiento.

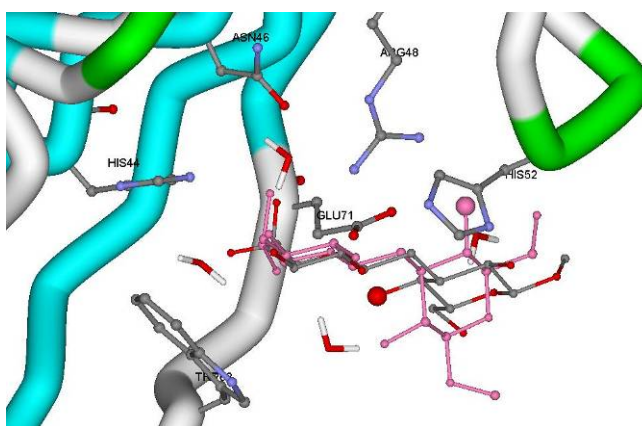


Figura 3.15: Superposición de C-glicósido en la conformación *syn-ψ(+)* (en rosa) y *syn-ψ(-)* (código de colores CPK). Se muestra la distinta orientación espacial de O4 (el oxígeno con tamaño agrandado) en ambas geometrías.

La otra conformación minoritaria en disolución, *anti-ψ*, también permite la interacción de las cadenas de Arg48 y Glu71 con un átomo de oxígeno del residuo Glc (**Figura 3.16** y **Figura 3.17**). En este caso, es el O2 Glc el que puede establecer enlaces de hidrógeno en lugar del O4 de la conformación *syn*. Por tanto, la menor energía estabilizante de esta geometría con respecto a la *syn-ψ(+)* (0.7 kcal mol⁻¹) puede ser debida a la existencia de menos interacciones entre ligando y receptor en el complejo con esta conformación, ya que se observa que una parte importante del residuo de Glc está en contacto con el medio acuoso (**Figura 3.18**)

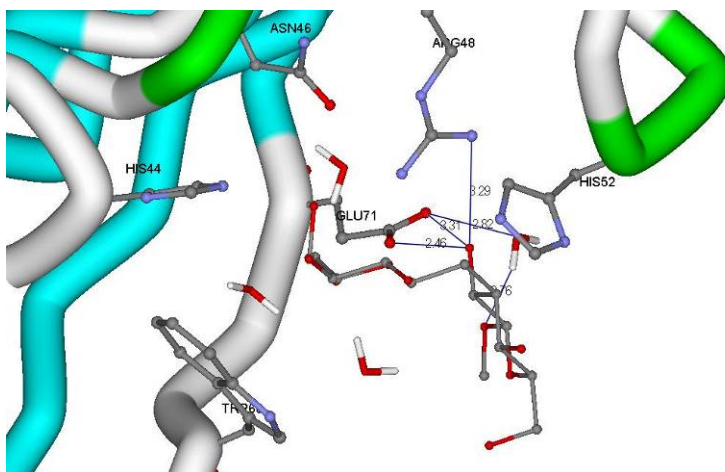


Figura 3.16: El C-glicósido en conformación *anti-ψ*. En este caso, la orientación espacial de O2 le permite formar enlaces de hidrógeno con las cadenas de Arg48 y Glu71.

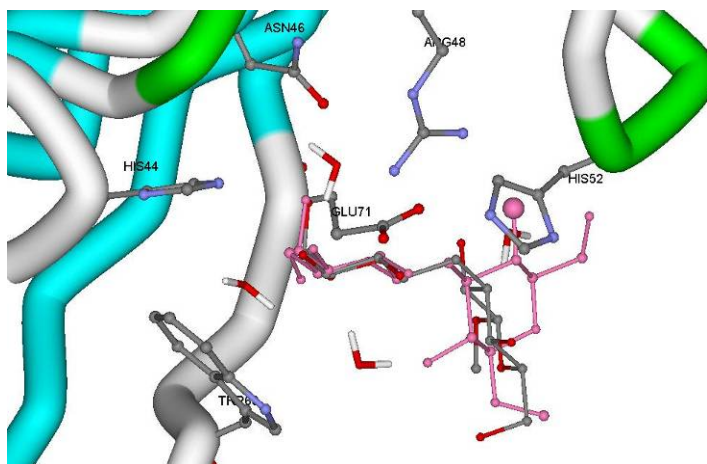


Figura 3.17: Superposición del C-glicósido en la conformación *syn-ψ(+)* (en rosa) y *anti-ψ* (código de colores CPK). La orientación del anillo de Glc es marcadamente diferente.

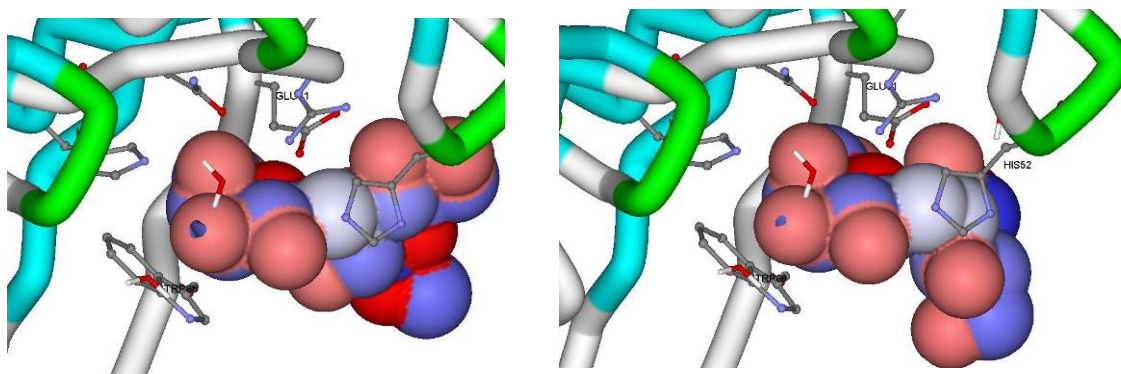


Figura 3.18: Superficies de van der Waals del C-glicósido en la conformación *syn-ψ(+)* (izquierda) y *anti-ψ* (derecha). La orientación del anillo de Glc es marcadamente diferente. En la conformación *anti* está más expuesta al disolvente y, por tanto, proporciona menos interacciones de van der Waals con la proteína.

5. Conclusiones:

El ensanchamiento de señales y los NOEs negativos que se observan para el disacárido **1** y el derivado C-glicosídico en presencia de la galectina-1 han demostrado que ambos compuestos son reconocidos por la proteína. Como los estudios de calorimetría y de RMN anteriores a este trabajo ya indicaban que la *N*-acetil-lactosamina se unía a la galectina-1, nuestros resultados demuestran que ni el *N*-acetilo ni el oxígeno interglicosídico son esenciales para el proceso de reconocimiento.

Además, los experimentos de NOE transferido indican que la conformación *syn*, que es la conformación reconocida por la galectina-1 para galactósidos $\beta(1\rightarrow4)$ es también la conformación bioactiva tanto en **1** como en **2**. Por tanto, el sitio de reconocimiento de la galectina-1 permite acomodar a la conformación *syn* independientemente del tipo de enlace entre los monosacáridos ($1\rightarrow4$ o $1\rightarrow3$) o la naturaleza química del puente (oxígeno o CH_2). Esta selección conformacional por parte de la galectina-1 ya se ha encontrado en otros complejos proteína-carbohidrato^{23,26-27}.

6. Parte Experimental:

6.1. Materiales:

La obtención de los compuestos **1** y **2** se describió en la parte experimental del capítulo 2. La galectina-1 fue proporcionada por el Prof. H. J. Gabius (Universidad LMU de Munich).

6.2. Estudio de Docking:

Preparación de la macromolécula: Se seleccionó la estructura cristalográfica de la galectina-1 humana complejada con la lactosa (código PDB: 1GZW). A continuación, se extrajo la molécula de lactosa y todas las moléculas de agua, a excepción de aquéllas que se consideraron esenciales para la unión de la lactosa a la proteína. Con ayuda del programa Sybyl 7.2³², se añadieron grupos neutros a los extremos *N*- y *C*-terminal. También se añadieron los hidrógenos polares, se orientaron las moléculas de agua de forma aleatoria, y se asignaron cargas puntuales de tipo Kollman United Atoms.

Preparación de los ligandos: Los ligandos estudiados fueron la lactosa, el compuesto **1** (*O*-disacárido) y el compuesto **2** (*C*-disacárido). La geometría inicial de la lactosa fue la que procedía de la estructura cristalográfica. La geometría inicial de **1** fue la correspondiente a la conformación de más baja energía obtenida en los cálculos de mecánica molecular (véase capítulo 2). Para el compuesto **2**, se consideraron geometrías iniciales diferentes correspondientes a las conformaciones de más baja energía obtenidas en los cálculos de mecánica molecular (véase capítulo 2). Con ayuda del programa Sybyl 7.2, se asignaron los tipos de átomos y los tipos de enlace. Se asignaron cargas del archivo de parámetros glycam04 de AMBER 8³³.

Estudio de acoplamiento o docking. El *docking* fue realizado con el programa AutoDock 3.1³¹, y el protocolo llevado a cabo y que se describe a continuación fue común para todos los ligandos. La región escogida se definió considerando una caja de 60, 60 y 60 puntos en los ejes *x*, *y*, y *z*. La caja fue construida centrándola en la posición que ocupa el oxígeno del enlace glicosídico de la lactosa en la estructura con código PDB 1GZW. Para el cálculo de los mapas de energía, se empleó un

espaciado de la malla de 0.375 Å y una función de la constante dieléctrica dependiente de la distancia. El *docking* se llevó a cabo empleando el algoritmo genético Lamarckiano. El programa buscó hasta un máximo de 100 conformaciones y el procedimiento fue repetido 50 veces. Para el resto de los parámetros se emplearon los valores que el programa proporciona por defecto. Las soluciones de *docking* generadas se agruparon en *clusters*. El análisis de los *clusters* se llevó a cabo empleando una tolerancia del RMSD (*root mean square deviation*) de 1.0 Å. Los *clusters* se colocaron por orden de energía (considerando la solución de menor energía dentro de cada cluster).

6.3. Experimentos de RMN:

Los espectros se adquirieron en un instrumento Bruker de 500 MHz, equipado con una criosonda inversa TCI. Las concentraciones de **1** y **2** fueron 18 y 1 mM, respectivamente. Todos los experimentos se llevaron a cabo a 295 K en D₂O, utilizando DSS como referencia externa. Los espectros en presencia de galectina-1 se llevaron a cabo en un instrumento Bruker de 500 MHz, equipado con una criosonda. Los experimentos tr-NOESY, tr-ROESY y tr-HSQC-NOESY se realizaron utilizando las secuencias habituales, con tiempos de mezcla de 150, 200 y 300 ms. Los mejores resultados se obtuvieron con una relación ligando:proteína de 20:1 en el *O*-Glicósido y de 5:1 en el *C*-análogo.

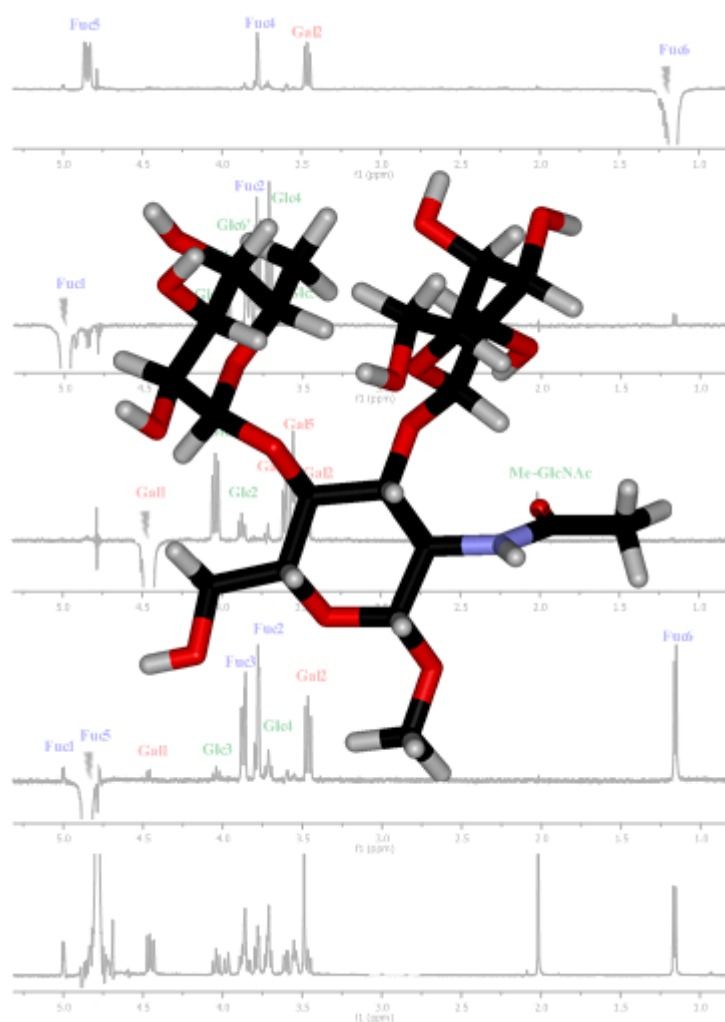
7. Bibliografía:

- (1) Poveda, A.; Jimenez-Barbero, J. *Chem Soc Rev* **1998**, 27, 133.
- (2) Hans-Joachim, G. *Eur J Biochem* **1997**, 243, 543.
- (3) Teichberg, V. I.; Silman, I.; Beitsch, D. D.; Resheff, G. *Proc Natl Acad Sci U S A* **1975**, 72, 1383.
- (4) de Waard, A.; Hickman, S.; Kornfeld, S. *J Biol Chem* **1976**, 251, 7581.
- (5) Cerra, R. F.; Gitt, M. A.; Barondes, S. H. *J Biol Chem* **1985**, 260, 10474.
- (6) Cummings, R. D.; Liu, F.-T. *Essentials of Glycobiology* 2nd Edition ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y. Atlanta, 2009.
- (7) Perillo, N. L.; Pace, K. E.; Seilhamer, J. J.; Baum, L. G. *Nature* **1995**, 378, 736.
- (8) Liu, F.-T.; Rabinovich, G. A. *Nat Rev Cancer* **2005**, 5, 29.
- (9) Camby, I.; Le Mercier, M.; Lefranc, F.; Kiss, R. *Glycobiology* **2006**, 16, 137R.
- (10) Camby, I.; Le Mercier, M.; Mathieu, V.; Ingrassia, L.; Lefranc, F.; Kiss, R. *Drugs Future* **2008**, 33, 1057.
- (11) Bourne, Y.; Bolgiano, B.; Liao, D. I.; Strecker, G.; Cantau, P.; Herzberg, O.; Feizi, T.; Cambillau, C. *Nat Struct Biol* **1994**, 1, 863.
- (12) Liao, D. I.; Kapadia, G.; Ahmed, H.; Vasta, G. R.; Herzberg, O. *Proc Natl Acad Sci U S A* **1994**, 91, 1428.
- (13) López-Lucendo, M. F.; Solís, D.; André, S.; Hirabayashi, J.; Kasai, K.-i.; Kaltner, H.; Gabius, H.-J.; Romero, A. *J Mol Biol* **2004**, 343, 957.
- (14) Mario, A. B.; Hafiz, A.; Gerardo, R. V.; Amzel, L. M. *Proteins: Struct, Funct, Genet* **2000**, 40, 378.
- (15) Nishi, N.; Abe, A.; Iwaki, J.; Yoshida, H.; Itoh, A.; Shoji, H.; Kamitori, S.; Hirabayashi, J.; Nakamura, T. *Glycobiology* **2008**, 18, 1065.
- (16) Brewer, C. F. *Glycoconj J* **2004**, 19, 459.
- (17) Meynier, C.; Feracci, M.; Espeli, M.; Chaspoul, F.; Gallice, P.; Schiff, C.; Guerlesquin, F.; Roche, P. *Biophys J* **2009**, 97, 3168.
- (18) Schwarz, F. P.; Ahmed, H.; Bianchet, M. A.; Amzel, L. M.; Vasta, G. R. *Biochemistry (Mosc)* **1998**, 37, 5867.
- (19) Ford, M. G.; Weimar, T.; Kohli, T.; Woods, R. J. *Proteins* **2003**, 53, 229.
- (20) Nesmelova, I. V.; Ermakova, E.; Daragan, V. A.; Pang, M.; Menéndez, M.; Lagartera, L.; Solís, D.; Baum, L. G.; Mayo, K. H. *J Mol Biol* **2010**, 397, 1209.
- (21) Siebert, H. C.; Andre, S.; Lu, S. Y.; Frank, M.; Kaltner, H.; van Kuik, J. A.; Korchagina, E. Y.; Bovin, N.; Tajkhorshid, E.; Kaptein, R.; Vliegenthart, J. F.; von der Lieth, C. W.; Jimenez-Barbero, J.; Kopitz, J.; Gabius, H. J. *Biochemistry (Mosc)* **2003**, 42, 14762.
- (22) Di Lella, S.; Ma, L.; Díaz Ricci, J. C.; Rabinovich, G. A.; Asher, S. A.; Álvarez, R. M. a. S. *Biochemistry (Mosc)* **2009**, 48, 786.
- (23) Alonso-Plaza, J. M.; Canales, M. A.; Jiménez, M.; Roldán, J. L.; García-Herrero, A.; Iturrino, L.; Asensio, J. L.; Cañada, F. J.; Romero, A.; Siebert, H.-C.; André, S.; Solís, D.; Gabius, H.-J.; Jiménez-Barbero, J. *Biochimica et Biophysica Acta (BBA) - General Subjects* **2001**, 1568, 225.
- (24) Nesmelova, I.; Pang, M.; Baum, L.; Mayo, K. *Biomol NMR Assignm* **2008**, 2, 203.
- (25) Westermann, J.-C., Universität Hamburg, 2006.
- (26) Asensio, J. L.; Espinosa, J. F.; Dietrich, H.; Canada, F. J.; Schmidt, R. R.; Martin-Lomas, M.; Andre, S.; Gabius, H. J.; Jimenez-Barbero, J. *J Am Chem Soc* **1999**, 121, 8995.

- (27) García-Aparicio, V.; Sollogoub, M.; Blériot, Y.; Colliou, V.; André, S.; Asensio, J. L.; Cañada, F. J.; Gabius, H.-J.; Sinaÿ, P.; Jiménez-Barbero, J. *Carbohydr Res* **2007**, *342*, 1918.
- (28) Ni, F. *Prog Nucl Magn Reson Spectrosc* **1994**, *26*, 517.
- (29) Post, C. B. *Curr Opin Struct Biol* **2003**, *13*, 581.
- (30) Moore, J. M. *Peptide Science* **1999**, *51*, 221.
- (31) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. *J Comput Chem* **1998**, *19*, 1639.
- (32) SYBYL; versión 7.2. ed.; Tripos: 1699 South Handley Road, St. Louis, Missouri 63144. , 1999, p <http://www.tripos.com>
- (33) D. A. Case, T. A. D., T. E. Cheatham, III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, K. M. Merz, B. Wang, D. A. Pearlman, M. Crowley, S. Brozell, V. Tsui, H. Gohlke, J. Mongan, V. Hornak, G. Cui, P. Beroza, C. Schafmeister, J. W. Caldwell, W. S. Ross, and P. A. Kollman; version 8 ed. University of California, San Francisco., 2004.

Capítulo 4

Estudio conformacional del trisacárido Lewis A: comparativa entre el *O*- y el *C*-trisacárido.



1. Introducción y antecedentes:

El Lewis a (Le^a) es uno de los principales carbohidratos antígenos que pertenece al sistema sanguíneo Lewis. El sistema Lewis está estructuralmente relacionado con el sistema ABO, que fue el primer grupo sanguíneo descubierto, por Landsteiner, en 1900¹. Los glóbulos rojos pueden clasificarse en A, B y O, de acuerdo con la presencia o ausencia de antígenos reactivos en su superficie. Dichos antígenos son claves para el éxito de la transfusión sanguínea, trasplante de tejidos y tienen un papel esencial en las enfermedades hemolíticas de recién nacidos.

El sistema Lewis es uno de los sistemas de eritrocitos inmunológicamente débiles. A diferencia del sistema ABO, no provoca reacciones inmunológicas severas. Los antígenos Lewis están formados por la adición secuencial de unidades de fucosa a las cadenas precursoras oligosacáridicas presentes en glicolípidos y glicoproteínas. Dependiendo de la estructura del disacárido al que se une la fucosa, se pueden clasificar en dos grupos: el tipo 1 contiene el disacárido β -Gal(1 \rightarrow 3)GlcNAc y a él pertenecen los antígenos Lewis a (Le^a) y Lewis b (Le^b). Por otra parte, el tipo 2 contiene el disacárido β -Gal(1 \rightarrow 4)GlcNAc, y a él pertenecen los antígenos Lewis y (Le^y) y Lewis x (Le^x) (**Figura 4.1**).

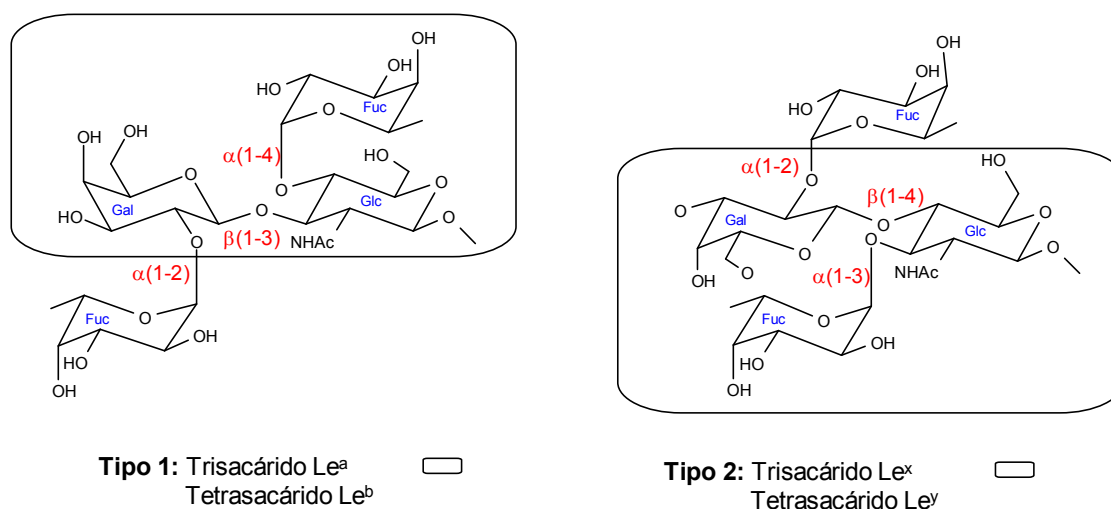


Figura 4.1: Representación esquemática de las estructuras Lewis de tipo 1 y de tipo 2.

A diferencia de otros antígenos de la sangre, los determinantes Le^y y Le^x sólo se expresan en algunos tipos de células, como las epiteliales. Por este motivo, al no estar asociados a los glóbulos rojos, no pueden definirse como verdaderos antígenos del grupo sanguíneo². Después se secretan al plasma como glicolípidos, para ser incorporados a las células eritrocíticas. Los antígenos del tipo 2 suelen expresarse, en personas sanas, en muy pequeña proporción. Sin embargo, se han observado en gran abundancia, asociados a la presencia de carcinomas (cáncer de colon, pulmón, hígado, próstata u ovario), por lo que se les considera como antígenos asociados a tumores, y podrían considerarse dianas para tratamientos de cáncer.

La función biológica de los antígenos Lewis parece ser la de actuar como ligandos de varias lectinas; en particular, las selectinas, ya que han mostrado ser activos en procesos de respuesta inflamatoria y de diferenciación³. Aunque su papel fisiológico no está todavía del todo claro, algunos compuestos sililados de tipo Lewis, así como otros análogos sulfatados pueden actuar como ligandos en procesos de adhesión celular.

Los primeros estudios estructurales de los determinantes Lewis fueron realizados por Lemieux⁴, a principios de los años 80. Los datos iniciales de RMN y de modelado molecular sugirieron que se trataba de estructuras bastante rígidas, y que poblaban una única familia conformacional en disolución. Este primer estudio se basó en un análisis comparativo entre los datos obtenidos para el trisacárido y los disacáridos que lo constituyen. Por ejemplo, se observó que el espectro del trisacárido Le^a era aproximadamente la suma de los espectros de los disacáridos constituyentes. Únicamente se observaron variaciones de desplazamiento químico en los protones Glc3 y Fuc5, que se atribuyeron al desapantallamiento provocado por la cercanía del protón H5 de la fucosa al O3 de la *N*-acetilglucosamina. De esta manera, se concluyó que la conformación en torno a los enlaces glicosídicos del Le^a se mantenía prácticamente inalterada con respecto a la existente en los disacáridos constituyentes. En los años 90, el laboratorio de Bush⁵ analizó los oligosacáridos de la leche humana, LNF-2 y LND-1, que contienen trisacáridos de Le^a y Le^b. A través de simulaciones de espectros NOESY, se encontró que únicamente un pequeño rango de ángulos diedros en torno a los enlaces glicosídicos, proporcionaban espectros NOESY que explicaban aquellos obtenidos experimentalmente, lo que ponía de

manifiesto su rigidez conformacional. Además de los estudios de RMN, Bush⁶ y sus colaboradores realizaron mapas conformacionales y simulaciones de dinámica molecular, que estaban de acuerdo con los datos experimentales. En concreto, se observó que tanto Le^a como Le^b eran conformacionalmente estables a lo largo de la simulación, obteniéndose un único mínimo conformacional. Estudios adicionales de estos antígenos, realizados por Kogelberg y Rutherford⁷, condujeron a conclusiones parecidas, en las que el Le^a fue descrito como una molécula rígida, (“*single-rigid molecule*”). Posteriormente, el grupo de Imberty⁸ realizó el análisis conformacional sistemático de todos los carbohidratos relacionados con los grupos sanguíneos, mediante cálculos de mecánica molecular, y compararon sus resultados con los parámetros experimentales de RMN y de Rayos-X, así como con los datos bioquímicos disponibles hasta ese momento. Los cálculos indicaron que los determinantes Lewis eran rígidos, aunque había que distinguir entre los carbohidratos de tipo 2 (Le^x y Le^y), que existían casi exclusivamente (>99%) en una conformación, denominada **I**, que se caracteriza por un apilamiento de los anillos de fucosa y galactosa, y los de tipo 1 (Le^a y Le^b), cuya conformación predominante era también la **I**, que representaba alrededor del 96% de la población total. Sin embargo, en éstos se predecía la coexistencia de otra familia conformacional minoritaria, **II**, con una pequeña población en torno al 4%. La conformación mayoritaria **I** es también el conformero más estable del Le^x en fase gaseosa, según los cálculos *ab initio* realizados por Csonk *et al.*⁹. Además, estudios conformacionales de los antígenos Le^a y Le^x mediante constantes dipolares residuales¹⁰, y de Le^b, Le^x y Le^y mediante el análisis de los protones hidroxílicos en agua¹¹, también estuvieron de acuerdo con el carácter monoconformacional de estos derivados. Posteriormente, Geyer *et al.*¹²⁻¹³ atribuyeron el carácter monoconformacional de los tetrasacáridos Le^b y Le^y a un efecto de doble apilamiento, “*two-fold stacking*”, definido por una disposición de los cuatro anillos de azúcar en un tetraedro ligeramente distorsionado, con el objeto de minimizar los contactos estéricos y de disponer las caras hidrófobas de la fucosa para interaccionar con el anillo de galactosa o de *N*-acetilglucosamina.

El objetivo de este capítulo ha sido realizar el estudio conformacional, en profundidad, del Le^a y de un mimético, en el que el oxígeno puente entre las unidades de galactosa y *N*-acetilglucosamina, se ha reemplazado por un grupo metilénico, para así determinar la influencia del cambio O→CH₂ en su

comportamiento conformacional. Como se ha explicado en detalle en la introducción del capítulo 2, los estudios previos de *O*- y *C*-glicósidos mediante RMN y cálculos de mecánica molecular han demostrado que esta sustitución produce cambios conformacionales en diferentes oligosacáridos, sobre todo alrededor del ángulo aglicónico, ψ . No obstante, la mayoría de los compuestos analizados hasta la fecha han sido disacáridos u oligosacáridos lineales, en los que el enlace pseudo-glicosídico no se encuentra en la vecindad de otro enlace glicosídico. En este tipo de oligosacáridos ramificados, con dos uniones glicosídicas contiguas, como los determinantes Lewis, se podría esperar que el cambio de un oxígeno por un CH₂ no produjera cambios conformacionales significativos, dada la menor libertad conformacional de estas moléculas. De hecho, un estudio previo de la conformación del antígeno Le^x y de un análogo *C*-glicosídico¹⁴, en el que el oxígeno puente entre las unidades de fucosa y glucosa fue sustituido por un grupo CH₂, reveló que ambas moléculas presentaban un comportamiento conformacional similar.

Aunque el trisacárido Le^a ha sido estudiado anteriormente por varios autores, decidimos realizar un nuevo estudio de RMN de este compuesto, utilizándose en este caso un espectrómetro equipado con una criosonda, que proporciona una sensibilidad mucho mayor que la de las sondas convencionales. Esto ha permitido la observación de NOEs débiles no detectados en los estudios previos, lo que ha llevado a una mejor caracterización de las propiedades conformacionales y dinámicas de este determinante sanguíneo.

2. Análisis conformacional del Lewis A:

El Le^a es el trisacárido β -Gal-(1 \rightarrow 3)-[α -Fuc-(1 \rightarrow 4)]- β -GlcNAc, donde Fuc, Gal y GlcNAc corresponden a *L*-Fucosa, *D*-Galactosa y 2-Acetamido-2-desoxi-*D*-Glucosa respectivamente. Se ha estudiado la mezcla anomérica (**3**) y su derivado *O*-metilado en la posición anomérica β (**3'**)

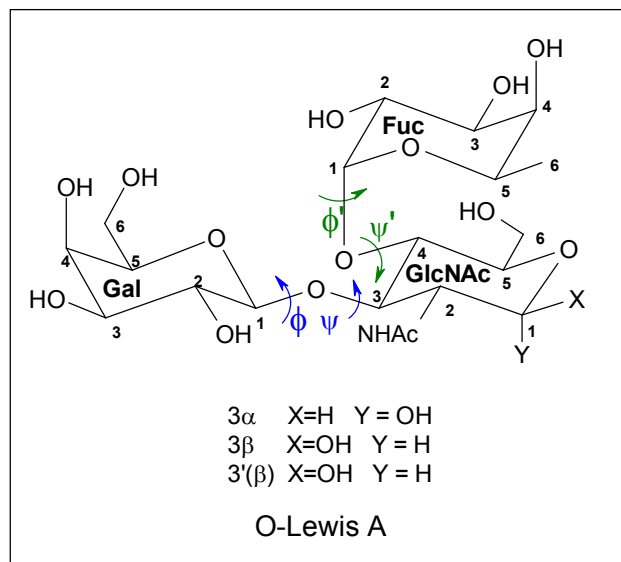


Figura 4.2: Estructura del Le^a, junto con la numeración y la designación de los ángulos diedros que serán utilizados en esta tesis.

2.1. Cálculos de Mecánica Molecular:

Como en los casos anteriores, el primer paso para el estudio de este compuesto fue la exploración de su espacio conformacional mediante cálculos de mecánica clásica. Sin embargo, la realización de una búsqueda sistemática, como en el caso de los disacáridos, daría lugar a cálculos largos y tediosos, ya que habría que explorar cuatro torsiones diferentes. El método alternativo elegido fue explorar el espacio conformacional mediante métodos estocásticos, en los que la búsqueda conformacional se realizó mediante la variación al azar de los ángulos interglicosídicos. La minimización y el cálculo de energía de las conformaciones resultantes se realizó mediante el campo de fuerzas MM3* ($\epsilon = 80$), como en los casos anteriores.

Estos cálculos indicaron la presencia de tres familias conformacionales diferentes con una población mayor del 0.1%, representadas por los mínimos **A-C**. Los mínimos **A** y **B** corresponden con los conformeros **I** y **II** previamente

identificados por Imberty y sus colaboradores⁸, aunque nosotros obtuvimos una población mayor para el conformero **II** (18% frente al 4%, ver **Tabla 4.1**). Los tres conformeros pertenecen a la misma región conformacional en torno a la unión Gal-GlcNAc, la región *syn-ψ(+)* siguiendo la nomenclatura del capítulo anterior. Las diferencias residen en la unión Fuc-GlcNAc: el conformero A se encuentra en la región *syn-ψ'(+)*, el conformero B en la región *syn-ψ'(-)*, y el conformero C en la región *non-exo'*. En la **Figura 4.3** se representan las distintas regiones conformacionales, en torno a cada unión interglucosídica, y la localización de los mínimos A(*), B(●) y C(■), mientras que en la **Figura 4.4** se detalla la geometría de los mismos.

		Le ^a (este estudio)			Le ^a (Imberty)	
		A (*)	B (●)	C (■)	I	II
Gal(1→3)GlcNAc	φ	53	75	58	53	76
	ψ	17	57	27	18	58
Fuc(1→4)GlcNAc	φ'	47	26	-21	47	28
	ψ'	23	-52	-38	23	-53
Energía (Kcal/mol)		0.0	0.9	2.8	0	1.8
Población		81	18	1	96	4
		<i>syn-ψ(+)</i>	<i>syn-ψ(+)</i>	<i>syn-ψ(+)</i>	<i>syn-ψ(+)</i>	<i>syn-ψ(+)</i>
		<i>syn-ψ'(+)</i>	<i>syn-ψ'(-)</i>	<i>non-exo'</i>	<i>syn-ψ'(+)</i>	<i>syn-ψ'(-)</i>

Tabla 4.1: Ángulos diedros, energías y poblaciones para los mínimos conformacionales del Le^a según este estudio y el realizado previamente por Imberty y colaboradores.

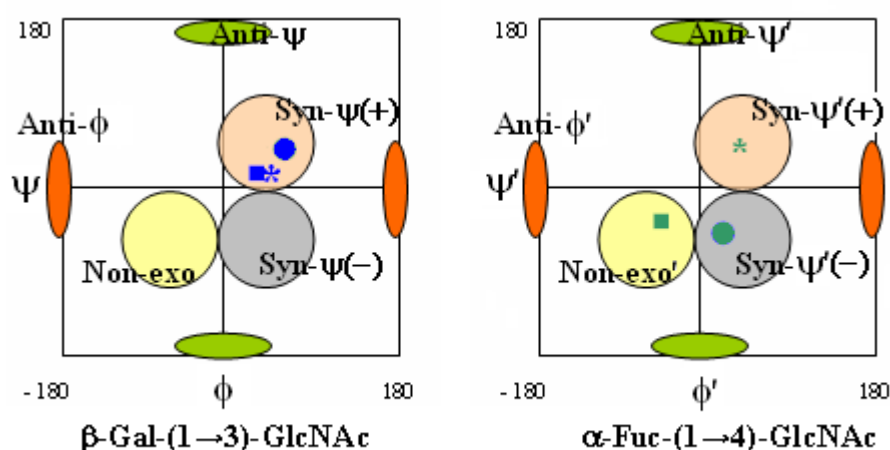


Figura 4.3: Representación sistemática de las diferentes regiones conformacionales y de la localización de los mínimos A (*), B (●) y C (■) del trisacárido Le^a.

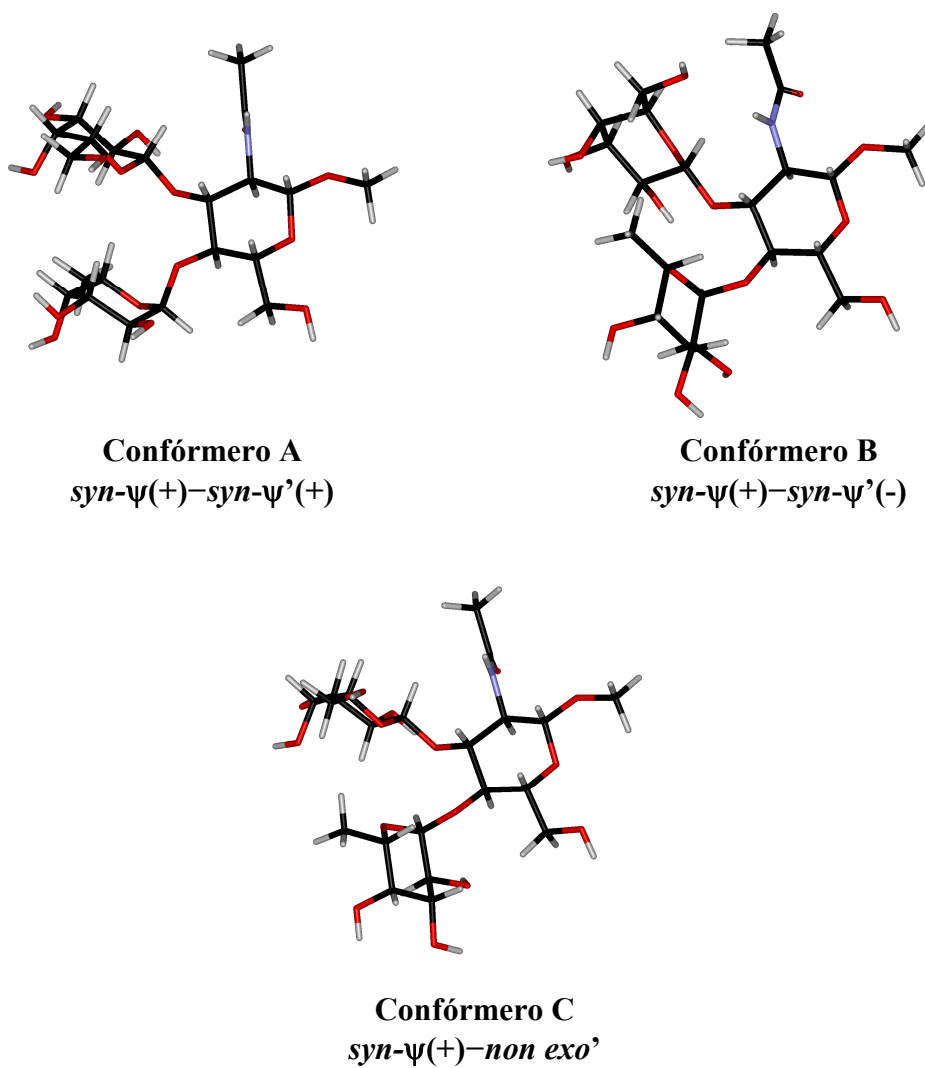


Figura 4.4: Mínimos conformaciones A, B y C para el Le^a, según el campo de fuerzas MM3* ($\epsilon = 80$). El primer término del nombre del confórmero se refiere a la unión Gal-GlcNAc; el segundo, a la unión Fuc-GlcNAc.

2.2. Estudio mediante RMN:

Siguiendo el protocolo descrito en el capítulo 1, se asignaron las resonancias de todos los protones de los anómeros α y β del trisacárido **3** y del derivado bloqueado en su forma anomérica β (**3'**). Para ello, se adquirieron experimentos de ^1H (**Figura 4.5**), COSY, HSQC, HMBC y 1D-TOCSY en D_2O . En la **Tabla 4.2** se recogen los desplazamientos químicos de protón de ambos compuestos.

3α	GlcNAc	Gal	Fuc	3β	GlcNAc	Gal	Fuc
H1	5.10	4.50	5.01	H1	4.70	4.46	5.00
H2	4.11	3.47	3.78	H2	3.85	3.47	3.78
H3	4.16	3.61	3.87	H3	4.03	3.61	3.87
H4	3.75	3.87	3.78	H4	3.72	3.86	3.78
H5	3.97	3.55	4.87	H5	3.54	3.55	4.87
H6α	3.94	3.70	–	H6α	3.97	3.71	–
H6β	3.90	3.70	–	H6β	3.87	3.71	–
Me(NAc)	2.02	–	–	Me(NAc)	2.02	–	–
Me(Fuc)	–	–	1.16	Me(Fuc)	–	–	1.16

3'(β)	GlcNAc	Gal	Fuc
H1	4.44	4.47	5.00
H2	3.88	3.46	3.78
H3	4.04	3.60	3.87
H4	3.71	3.86	3.78
H5	3.54	3.55	4.86
H6a	3.98	3.71	–
H6b	3.88	3.71	–
Me(NAc)	2.01		
Me(OMe)	3.49	–	–
Me(Fuc)	–	–	1.16

Tabla 4.2: Desplazamientos químicos de protón para **3** y **3'** en D_2O a 25°C .

Para la detección de NOEs, se realizaron experimentos 2D-NOESY, 2D-ROESY y varios experimentos 1D-NOESY (**Figura 4.5**), donde se excitaron selectivamente los protones anoméricos, Fuc1 y Gal1, el metilo de la fucosa (Fuc6) y el protón 5 de la fucosa (Fuc5).

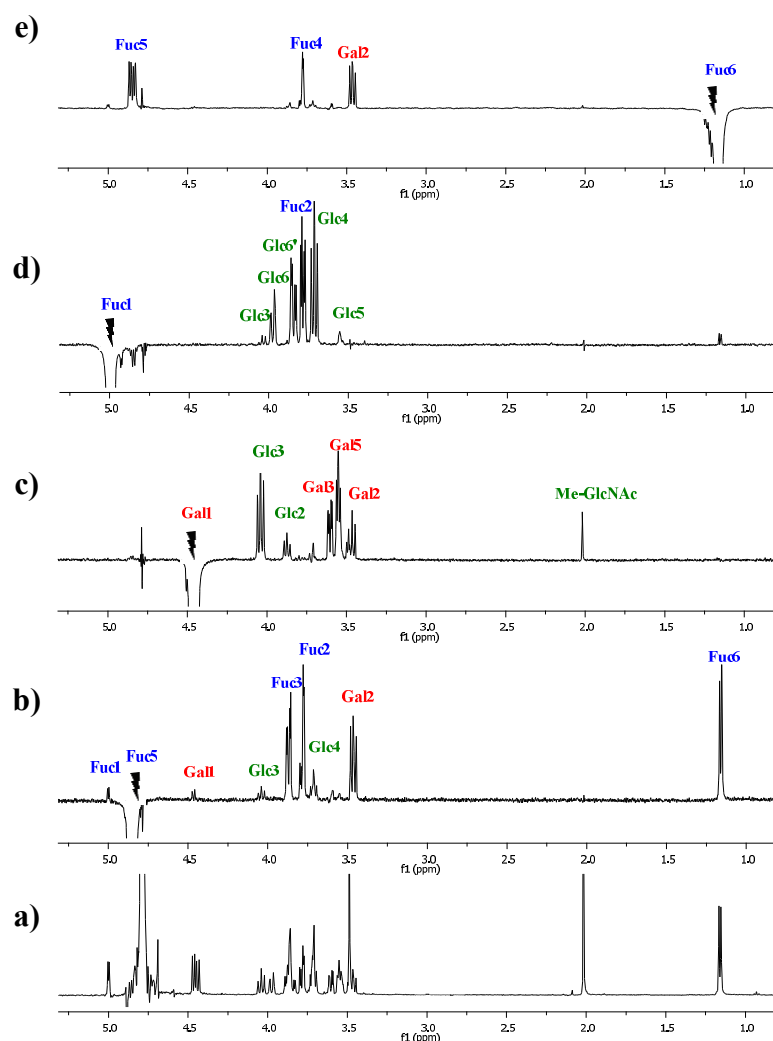


Figura 4.5: Espectros de protón (a), y 1D-NOESY (b-e) con tiempo de mezcla de 500 ms, del trisacárido **3'**.

Los NOEs observados en los compuestos **3** y **3'** fueron los mismos, confirmando que la presencia de un metilo, en la posición de la GlcNAc, no afecta a la conformación del trisacárido. Dado que la cuantificación de los NOEs es más sencilla en el anómero bloqueado **3'**, en el que hay un menor número de señales, se utilizó este compuesto para el estudio cuantitativo.

En la **Tabla 4.3** se han recogido los NOEs observados en este estudio clasificados en intensos (i), medios (m) y débiles (d), según su intensidad. Además, en esa misma tabla se recogen los NOEs observados en distintos estudios previos del Le^a y de oligosacáridos que contienen este trisacárido. Como se puede apreciar, los NOEs clasificados como intensos fueron detectados en estudios anteriores. Sin embargo, los catalogados como débiles o de intensidad intermedia sólo han sido detectados en este estudio.

		NOEs	LNF2 ⁵	Le ^{a7}
Fuc1	GlcNAc3	d		
	GlcNAc4	i	x	x
	GlcNAc5	d		
	GlcNAc6	i		x
	GlcNAc6'	i		
Gal1	GlcNAc2	m		
	GlcNAc3	i	x	x
Fuc5	Gal1	d		
	Gal2	i	x	x
	GlcNAc3	d		
	GlcNAc4	x		
Fuc6	Gal2	i		x

Tabla 4. 3: Tabla comparativa de los NOEs observados en este estudio y en otros anteriores, indicándose las referencias bibliográficas en la primera fila. Los NOEs de este análisis se clasifican en intensos (i), medios (m) y débiles (d), salvo cuando su intensidad no se pudo cuantificar debido al solapamiento. El LNF2 es un oligosacárido de la leche llamado Lacto-N-fucopenteosa, que contiene al trisacárido Le^a.

De nuevo, para sacar conclusiones sobre el comportamiento conformacional de este compuesto a partir de los datos de NOE, es necesario apoyarse en los resultados computacionales previos. Como se ha mencionado anteriormente, los seis NOEs más intensos, Fuc1-GlcNAc4, Fuc1-GlcNAc6, Fuc1-GlcNAc6', Gal1-GlcNAc3, Fuc6-Gal2 y Fuc5-Gal2, fueron ya observados en estudios anteriores y son NOEs esperados para el mínimo global **A** (**I** según la nomenclatura de Imberty), tal y como se deduce de las distancias interprotónicas recopiladas en la **Tabla 4.4**.

		Conf.A	Conf.B	Conf.C
Fuc1	GlcNAc3	4.5	3.6	3.5
	GlcNAc4	2.5	2.3	2.4
	GlcNAc5	3.7	4.5	4.5
	GlcNAc6	2.5	4.4	4.6
	GlcNAc6'	2.1	4.0	4.3
	Fuc5	3.7	3.7	3.7
Gal1	GlcNAc2	3.7	2.8	3.5
	GlcNAc3	2.5	3.3	2.7
	Gal3	2.7	2.6	2.6
Fuc5	Gal1	3.9	5.5	6.0
	Gal2	2.4	4.6	5.7
	GlcNAc3	3.7	2.7	4.0
	GlcNAc4	3.3	4.4	4.5
	Fuc1	3.7	3.7	3.7
Fuc6	Gal2	3.1	4.0	5.9

Tabla 4.4: Distancias interprotónicas relevantes para cada uno de los conformeros. Las distancias menores de 3 Å están marcadas en azul.

En concreto, uno de ellos, el NOE Fuc5-Gal2, puede considerarse como exclusivo de esta conformación ya que estos protones están a más de 4.5 Å de distancia en los otros dos conformeros posibles (**B** y **C**), lo que demuestra la presencia del conformero **A** en disolución. Este resultado está de acuerdo con los estudios anteriores que pusieron de manifiesto que el Le^a adoptaba la conformación **A** en disolución acuosa. La observación de otros NOEs adicionales de intensidad débil (Fuc1-GlcNAc5, Fuc5-Gal1 y Fuc5-GlcNAc3), que no habían sido detectados anteriormente, también es consistente con este resultado, ya que son NOEs entre protones separados entre 3.5 y 4 Å en el conformero **A**. Su observación en este estudio podría atribuirse a la mayor sensibilidad que proporciona la criosonda utilizada, que permite detectar y cuantificar, con exactitud, NOEs de alrededor del 0.1%. Sin embargo, la observación del NOE Fuc1-GlcNAc3, entre protones que se encuentran a más de 4.5 Å de distancia en el conformero **A**, junto al hecho de que el NOE Gal1-GlcNAc2 es mucho más intenso que el esperado para el conformero **A**, sugiere también la presencia del conformero **B** (**II** según la nomenclatura de Imberty) en disolución. En la **Tabla 4.5** se recoge la intensidad de los NOEs experimentales junto a los NOEs calculados mediante la aproximación de la matriz de relajación completa¹⁵.

		NOEs Teóricos				NOEs Experimentales
		Conf.A	Conf.B	Conf.C	Distribución	%NOE
Fuc1	GlcNAc3	-	0.2	0.3	0.1	0.2
	GlcNAc4	2.2	3.1	2.6	2.4	2.6
	GlcNAc5	0.2	-	0.1	0.1	0.3
	GlcNAc6	1.4	-	-	1.1	1.1
	GlcNAc6'	5.1	0.1	0.1	4.2	2.0
	Fuc2	2.7	2.3	2.3	2.6	2.6
Gal1	GlcNAc2	0.2	1.0	0.3	0.3	0.6
	GlcNAc3	2.1	0.4	1.3	1.8	2.4
	Gal3	1.5	1.5	1.5	1.5	1.4
Fuc5	Gal1	0.1	-	-	0.1	0.2
	Gal2	2.4	-	-	2.0	2.3
	GlcNAc3	0.2	1.2	0.1	0.4	0.3
	GlcNAc4	0.4	0.1	0.1	0.3	No cuantificado
	Fuc3	1.7	1.6	1.5	1.7	2.7
Fuc6	Gal2	1.2	-	-	1.0	1.3
	Fuc5	1.9	1.8	1.8	1.8	1.9

Tabla 4.5: NOEs teóricos, obtenidos a partir de la matriz de relajación completa, frente a NOEs experimentales, medidos en experimentos 1D-NOESY, con 500 ms de tiempo de mezcla. Los NOEs intra-residuo, Fuc1-Fuc5 y Gal1-Gal3, fueron utilizados como referencia para la estimación del tiempo de correlación efectivo (38 ps).

Una vez que se ha concluido que para justificar los datos de NOE es necesaria la presencia de los confórmeros **A** y **B** en disolución, se procedió a estimar la población de cada uno de ellos mediante un parámetro, conocido como factor R, cuya fórmula es:

$$R = \sum \frac{|NOE_{exp.} - NOE_{teo.}|}{NOE_{exp.}}$$

Este parámetro ha sido propuesto anteriormente para determinar el grado de ajuste entre NOEs experimentales y calculados para la estructura de una biomolécula¹⁶, y se puede utilizar para determinar la proporción de confórmeros cuyos NOEs teóricos mejor se ajustan a los datos experimentales. Para ello se calcularon los NOEs para diferentes proporciones de **A** y **B**, obteniéndose el valor más pequeño para el factor R con una proporción **A:B** de 95:5, como se representa en la **Figura 4.6**. Por tanto, los cálculos teóricos de MM3* sobrestiman la población del confórmero minoritario **B**, para el que se predijo una población del 18%.

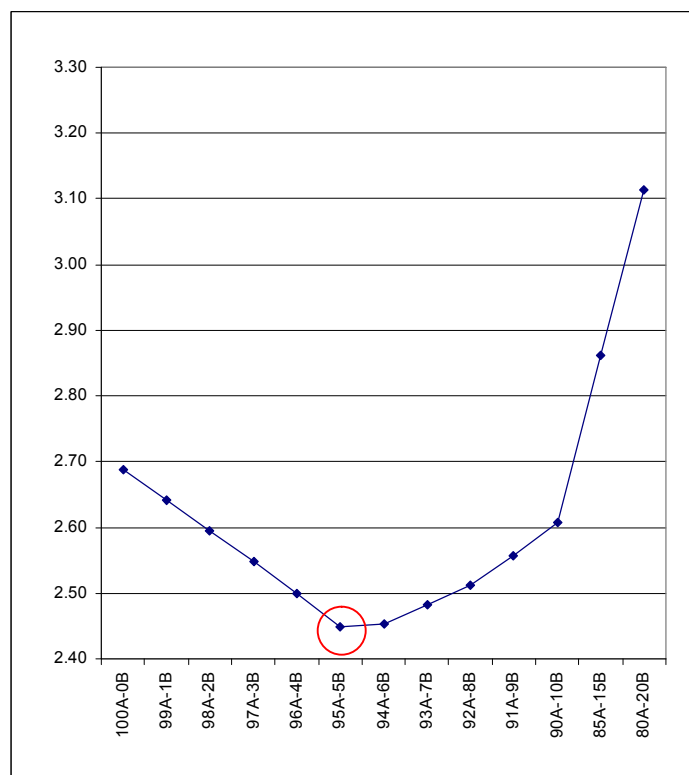


Figura 4.6: Representación del error para cada proporción de confórmeros.

2.3. Conclusiones:

Al contrario que en estudios anteriores en los que mediante RMN, sólo se había podido demostrar la existencia del confórmero **A** en disolución, en este estudio se ha visto que, aunque esta conformación es muy predominante en disolución, los datos de NOE se explican mejor con la presencia en disolución de un segundo confórmero (**B**), que se diferencia del **A** en el valor de la torsión ψ' de la unión Fuc-GlcNAc (*syn- ψ'* (-) frente a *syn- ψ'* (+)), aunque con una población pequeña, en torno al 5%.

Este estudio revela que la flexibilidad del antígeno Le^a no sólo se limita a los movimientos internos detectados previamente mediante estudios de relajación en otros determinantes Lewis¹⁷, sino a la coexistencia de dos familias conformacionales, que se diferencian en la orientación del enlace aglicónico de la unión Fuc α (1 \rightarrow 4)Glc. Por tanto, la conclusión previa acerca del carácter monoconformacional de este compuesto no es del todo correcta. Es interesante resaltar que este resultado confirma las predicciones computacionales de Imberty y colaboradores que anticiparon la presencia de los confórmeros A:B en proporción 96:4, en excelente acuerdo con nuestros datos experimentales.

3. Estudio del análogo modificado, C-Lewis A:

El C-Le^a (**4**) corresponde al análogo trisacárido Le^a, donde el oxígeno interglucosídico entre las unidades de galactosa y *N*-acetilglucosamina ha sido sustituido por un grupo metilénico (**Figura 4.7**).

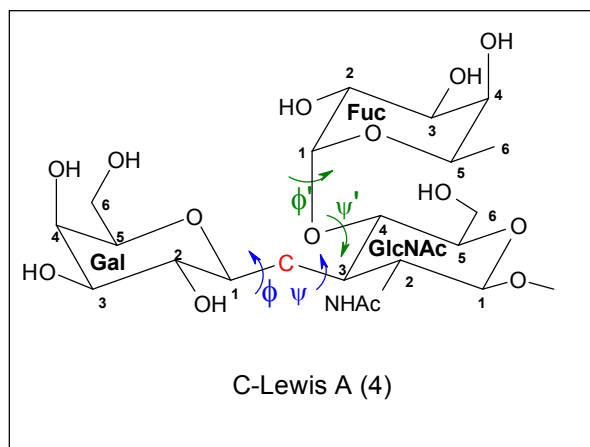


Figura 4.7: Estructura del C-Le^a junto con la numeración y la designación de los ángulos diedros que serán utilizados en esta tesis.

3.1. Cálculos de Mecánica Molecular:

Como en el caso del trisacárido natural, se realizó una búsqueda conformacional estocástica, siguiendo el mismo procedimiento empleado anteriormente. Para facilitar la comparación de ambos compuestos, se utilizarán las letras **A**, **B** y **C** para designar los confórmeros situados en las mismas regiones conformacionales que en el caso del trisacárido natural.

Los cálculos predijeron un comportamiento conformacional distinto al del análogo natural, con la presencia de una familia conformacional mayoritaria (confórmero **A**) y otras dos familias conformacionales sustancialmente pobladas, confórmeros **D** y **E**. Además de estos tres mínimos, se predicen otras cinco familias conformacionales, con una población entre el 1 y el 5%. En la **Tabla 4.6** se han recogido los ángulos diedros y energías relativas de cada uno de los confórmeros, mientras en la **Figura 4.8** y **4.9** se han representado las distintas regiones conformacionales y los tres mínimos mayoritarios, **A**(*), **D**(•) y **E**(■).

C-Le ^a (4)									
	A (*)	D (●)	E (■)	F	B	G	H	I	
Gal(1→3)GlcNAc	55	63	47	41	68	-46	53	54	ϕ
	16	42	174	-54	50	-22	-172	-172	ψ
Fuc(1→4)GlcNAc	34	26	72	56	18	52	20	28	ϕ'
	23	173	56	42	-37	37	-29	168	ψ'
Energía (Kcal/mol)	0.0	0.7	0.9	1.6	1.8	2.1	2.2	2.3	
Población	57	18	14	4	3	2	1	1	
Gal(1→3)GlcNAc	syn-ψ(+)	syn-ψ(+)	anti-ψ	syn-ψ(-)	syn-ψ(+)	non exo	anti-ψ	anti-ψ	
Fuc(1→4)GlcNAc	syn-ψ'(+)	anti-ψ'	syn-ψ'(+)	syn-ψ'(+)	syn-ψ'(-)	syn-ψ'(+)	syn-ψ'(-)	anti-ψ'	

Tabla 4.6: Ángulos diedros, energías y poblaciones para los mínimos del compuesto 4, C-Le^a, según los cálculos de mecánica molecular con el campo de fuerzas MM3* (ε = 80).

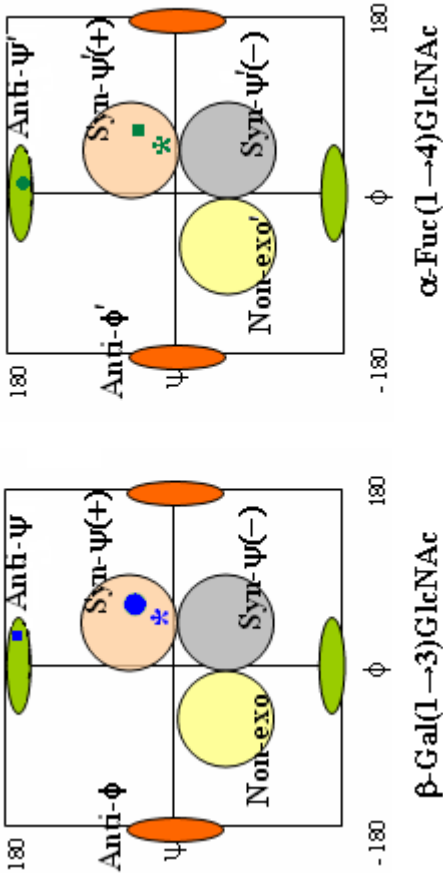


Figura 4.8: Representación sistemática de las diferentes regiones y de los mínimos principales A (*), D (●) y E (■).

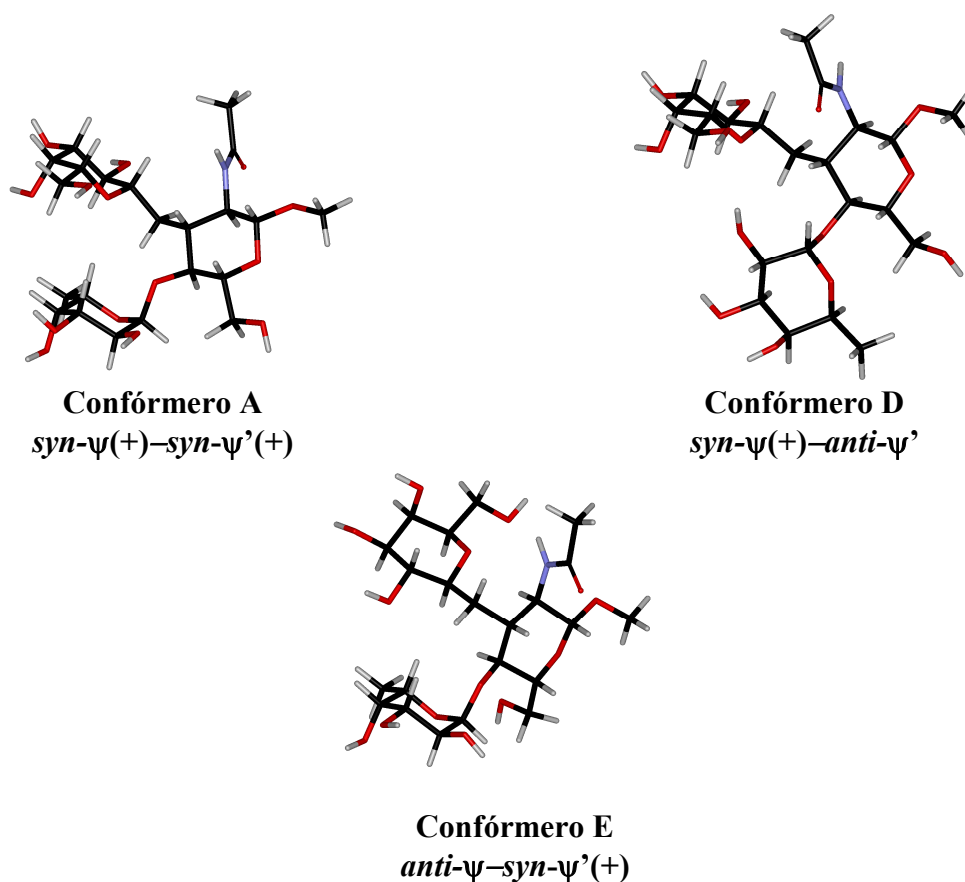


Figura 4.9: Mínimos A, D y E para el C-Le^a.

Si se comparan los resultados de mecánica molecular del Le^a y del C-Le^a, se observa que, en ambos casos, el mínimo global se encuentra en la misma región conformacional, (*syn-ψ(+)-syn-ψ'(+)*). Sin embargo, mientras que en el análogo natural, la mayor parte de la población reside en esta región (81% según los cálculos de mecánica molecular), la población de **A** se reduce al 57% en el C-análogo. Por otro lado, en el Le^a, se predice un segundo confórmero, **B**, localizado en la región *syn-ψ(+)-syn-ψ'(-)*, con una población del 18%. Este confórmero **B** apenas está poblado en el C-trisacárido (3%), mientras que dos regiones que no estaban pobladas en el trisacárido natural, las correspondientes a las familias conformacionales *syn-ψ(+)-anti-ψ'* y *anti-ψ-syn-ψ'(+)*, representadas por los mínimos **D** y **E**, sí están pobladas (18 y 14% de la población total respectivamente) en el glicomimético. En general, los cálculos de mecánica molecular sugieren una mayor flexibilidad conformacional en torno a las uniones interglucosídicas con respecto al trisacárido natural, con la existencia de confórmeros *anti* en torno a ambos enlaces aglicónicos.

3.2. Estudio mediante RMN:

De nuevo, el primer paso fue la asignación de todos los protones, mediante la combinación de experimentos de ^1H , COSY, HSQC, HMBC y 1D-TOCSY. En la **Tabla 4.7** se recogen los desplazamientos químicos de protón del anómero β , que es el mayoritario en disolución acuosa.

β anómero	GlcNAc	Gal	Fuc
H1	4.38	3.15	5.02
H2	3.62	3.32	3.72
H3	2.15	3.47	3.69
H4	3.48	3.91	3.71
H5	3.52	3.47	4.48
H6a	4.03	3.67	—
H6b	3.90	3.67	—
Me (NAc)	2.01	—	—
7pR	—	1.72	—
7pS	—	1.90	—
Me (Fuc)	—	—	1.19
Me (OMe)	3.47	—	—

Tabla 4.7: Desplazamientos químicos de protón (en ppm) de **4** en D_2O a 25°C .

Como es habitual en *C*-glicósidos, en el *C*-Le^a se observa mayor dispersión de señales que en el trisacárido natural, apareciendo la mayoría de las señales claves para el estudio conformacional bien separadas del resto, lo que facilitó el análisis de los espectros.

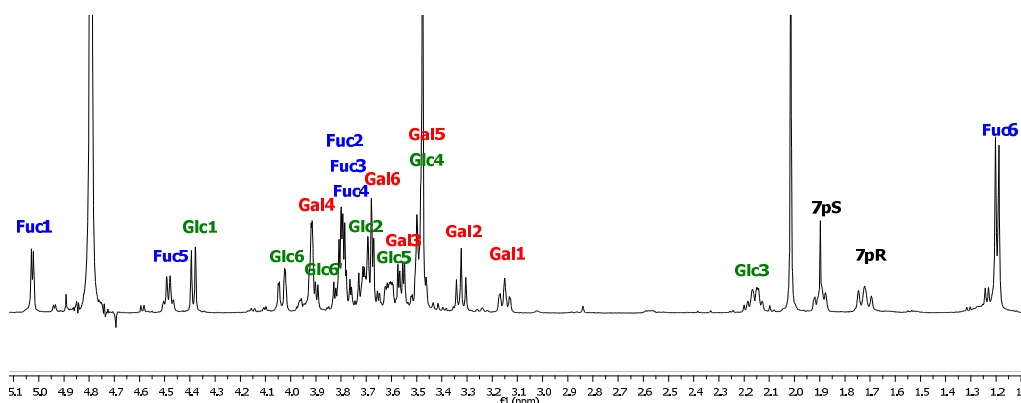


Figura 4.10: Espectro de protón del *C*-Le^a en el que se indica la asignación de las señales.

Como en el caso del *C*-disacárido **2**, la presencia de dos protones en el carbono que conecta las unidades de Gal y GlcNAc permitió medir las constantes de

acoplamiento homonucleares a través de esta unión, lo que proporcionó información conformacional complementaria a la que se obtiene mediante el análisis de NOEs. Para la obtención de valores más exactos, se realizó un análisis de segundo orden obteniéndose un buen ajuste entre el espectro experimental y el teórico (**Figura 4.11**). Los valores determinados experimentalmente se compararon con los calculados para los mínimos **A-I** y para una distribución de Boltzmann, estimada a partir de los valores relativos de energía MM3*, según la ecuación de Karplus-Altona (**Tabla 4.8**).

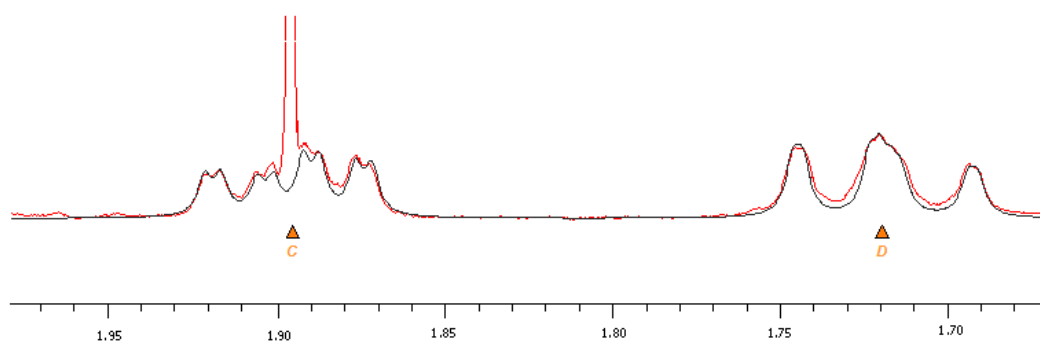


Figura 4.11: Ampliación de la zona del espectro de protón de **4** en la que aparecen los protones metilénicos (en rojo) superpuesta al espectro simulado (en negro).

$^3J_{HH}$		A	D	E	F	B	G	H	I	Exp.	Distribución
Gal1	7pR	11.4	11.7	10.9	10.5	11.6	1.9	11.4	11.5	11.5	10.9
Gal1	7pS	1.5	2.1	1.0	1.0	2.7	11.5	1.3	1.4	2.4	1.6
GlcNAc3	7pR	2.0	1.2	4.4	12.3	1.8	8.8	2.5	2.6	1.5	2.8
GlcNAc3	7pS	7.5	11.7	2.6	2.3	12.2	1.4	4.8	4.5	7.9	7.1
Población		57	18	14	4	3	2	1	1		
		syn- $\psi(+)$	syn- $\psi(+)$	anti- ψ	syn- $\psi(-)$	syn- $\psi(+)$	non exo	anti- ψ	anti- ψ		

Tabla 4.8: Constantes de acoplamiento (en Hz) a través del enlace pseudo-glicosídico para los mínimos **A-I** de **4** y para la distribución teórica junto con los valores experimentales en D₂O a 25°C.

Como puede observarse, existe un magnífico ajuste entre las constantes de acoplamiento determinadas para la distribución y las medidas experimentalmente, lo que, en principio, apoya la validez de los cálculos de mecánica molecular. Sin embargo, es importante señalar que las constantes esperadas para el mínimo global **A** son también muy similares a las calculadas de la distribución, por lo que los valores experimentales también cuadrarían con la presencia de una única familia conformacional en disolución, en torno al enlace Gal→GlcNAc. Por tanto, las constantes de acoplamiento no permiten distinguir entre estas dos situaciones, y es

necesario recurrir al análisis de los NOEs para caracterizar el comportamiento conformacional del C-Le^a. No obstante, la comparación entre las constantes de acoplamiento experimentales y teóricas sí sirvió para confirmar que las familias conformacionales *syn-ψ(-)* y *non-exo*, representadas por los mínimos **F** y **G**, respectivamente, apenas se encuentran poblados en disolución, ya que estos confórmeros disponen los protones GlcNAc3 y H7_{pR} (mínimo **F**), o los protones Gal1 y H7_{pS} (mínimo **G**), en disposición *anti*-periplanar, por lo que si estuvieran significativamente poblados en disolución, las constantes de acoplamiento deberían ser mayores que los valores experimentales de 1.5 y 2.4 Hz, respectivamente.

A continuación, se adquirieron distintos experimentos 2D-NOESY y 1D-NOESY. En éstos últimos, se excitaron selectivamente los protones anoméricos, Fuc1 y Gal1, el metilo de la fucosa (Fuc6) y el protón 5 de la fucosa (Fuc5). Los NOEs observados (ver **Figura 4.12**) se recogen en la **Tabla 4.9**, junto a los NOEs previamente observados para el Le^a, con el objeto de facilitar la comparación entre ambos compuestos. Se clasificaron en intensos, medios y débiles, como en el caso anterior.

		Le ^a		C-Le ^a	
		NOE	Intensidad	NOE	Intensidad
Fuc1	GlcNAc3	0.2	d	0.6	m
	GlcNAc4	2.6	i	2.6	i
	GlcNAc5	0.3	d	0.4	d
	GlcNAc6	1.1	i	1.0	i
	GlcNAc6'	2.0	i	1.3	i
	7pR	-	-	0.2	d
Gal1	GlcNAc2	0.6	m	0.4	d
	GlcNAc3	2.4	i	1.2	i
	GlcNAc4	-	-	-	-
	7pR	-	-	0.6	m
	7pS	-	-	1.3	i
	Gal3	1.4	i	1.4	i
Fuc5	Gal1	0.1	d	0.2	d
	Gal2	2.4	i	0.7	m
	GlcNAc3	0.2	d	0.3	d
	GlcNAc6	NO	-	0.3	d
	GlcNAc6'	NO	-	0.3	d
	7pR	-	-	2.2	i
	Fuc1	0.2	d	0.3	d
Fuc6	Gal2	1.2	i	0.7	m
	7pR	-	-	0.3	d

Tabla 4.9: Tabla comparativa de los NOEs encontrados para el O-Lewis a y para su C-derivado. Los NOEs fueron clasificados en intensos(i), medios(m) y débiles(d) según su intensidad, medida en experimentos 1D-NOESY, con un tiempo de mezcla de 500ms.

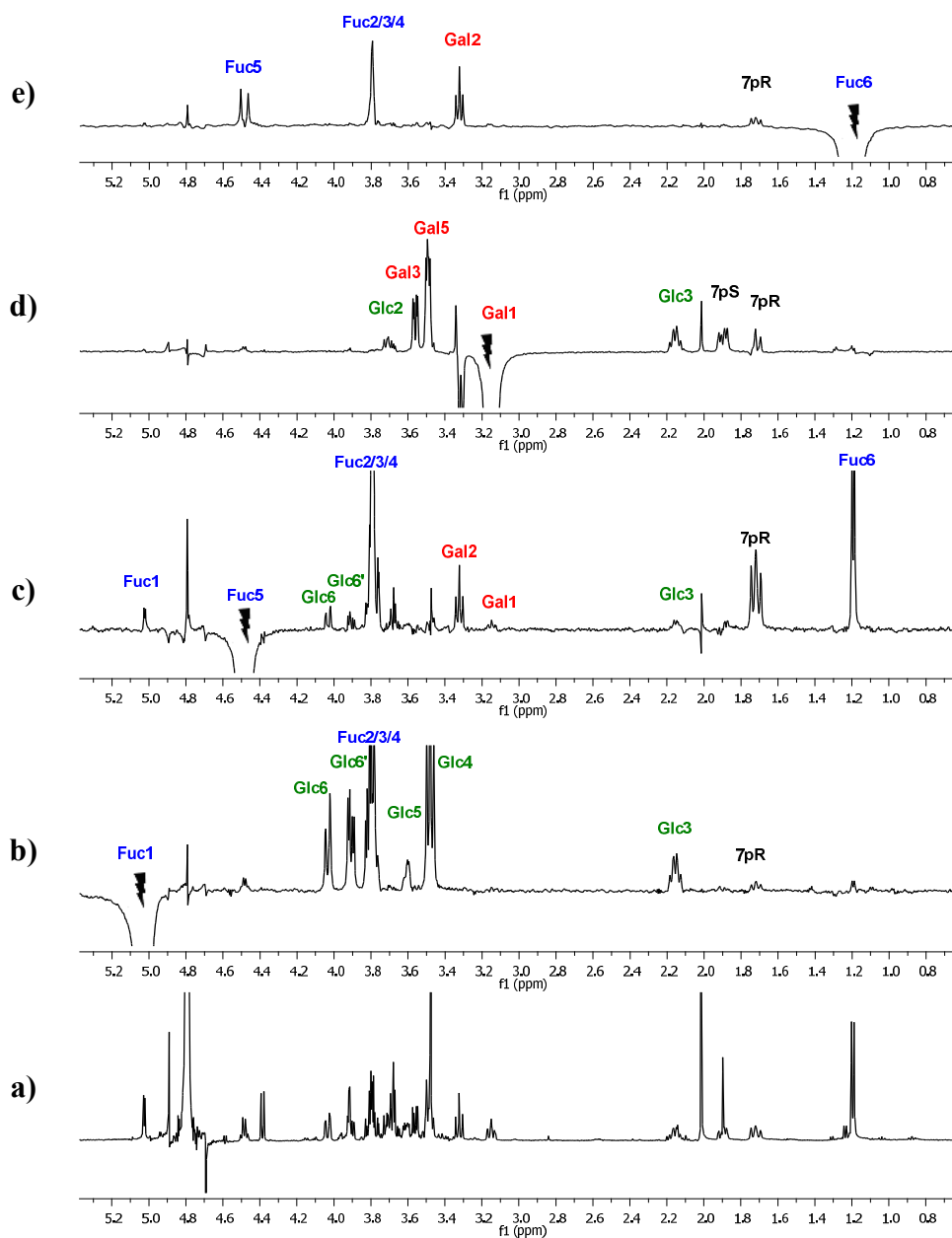


Figura 4.12: Espectros de protón (a), y 1D-NOESY (b-e) con 500ms de tiempo de mezcla del C-trisacárido 4.

Como se describió anteriormente, en el trisacárido Le^a se observaron seis NOEs (Fuc1-GlcNAc4, Fuc1-GlcNAc6, Fuc1-GlcNAc6', Gal1-GlcNAc3, Fuc5-Gal2 y Fuc6-Gal2), significativamente más intensos que el resto de NOEs inter-residuo, debido a la cercanía de los protones correspondientes en el conformero mayoritario (95%), *syn-ψ(+)-syn-ψ'(+)* (conformación A). Estos NOEs siguen siendo de intensidad fuerte o media en el C-Le^a, lo que sugiere que esta familia conformacional continúa dominando el equilibrio conformacional del análogo C-glicósido, tal y

como predicen los cálculos teóricos. Además, la presencia de dos protones metilénicos en el carbono puente permite la observación de dos NOEs intensos adicionales, Fuc5-H7pR y Gal1-H7pS, que también son consistentes con el predominio de este conformero en disolución.

Sin embargo, se detectaron otros NOEs (Fuc1-GlcNAc3, Fuc5-GlcNAc6 y Fuc5-GlcNAc6') que no pueden justificarse con el conformero *syn-ψ'*(+) en torno a la unión Fuc-GlcNAc, ya que los protones correspondientes se encuentran demasiado alejados en este conformero (distancia superior a 4.5 Å) como para que exista un NOE observable entre ellos. Este hecho pone de manifiesto que al menos una segunda familia conformacional debe estar poblada en disolución. Estos NOEs pueden considerarse exclusivos del conformero *anti-ψ'*, representado por el mínimo D, como se deduce de las distancias interprotónicas recopiladas en la **Tabla 4.10**. Por tanto, la observación de estos NOEs confirma la presencia de la familia conformacional *anti-ψ'* en disolución acuosa, tal y como predicen los cálculos computacionales.

		Exp NOEs	A	D	E	B	H	I
Fuc1	GlcNAc3	m	4.5	2.0	4.2	3.9	4.1	2.1
	GlcNAc4	i	2.4	3.7	3.3	2.3	2.2	3.7
	GlcNAc5	d	3.7	2.3	2.9	4.5	4.4	2.3
	GlcNAc6	i	2.1	3.8	3.6	3.8	3.6	3.8
	GlcNAc6'	i	2.7	4.8	1.9	4.1	3.9	4.8
	7pR	d	3.8	3.5	5.7	2.3	4.5	4.3
Gal1	GlcNAc2	d	3.9	3.4	2.8	3.1	3.2	3.3
	GlcNAc3	i	2.6	3.1	3.9	3.3	3.8	3.8
	GlcNAc4	NO	4.7	4.6	2.2	4.6	2.2	2.1
	7pR	m	3.1	3.1	3.1	3.1	3.1	3.1
	7pS	i	2.6	2.5	2.6	2.5	2.6	2.5
Fuc5	Gal1	d	4.2	7.5	2.3	5.5	4.6	4.9
	Gal2	m	3.0	6.9	4.3	4.9	5.6	7.4
	GlcNAc3	d	3.7	5.0	4.3	2.9	3.0	5.0
	GlcNAc6	d	5.2	2.2	5.7	4.6	4.7	2.2
	GlcNAc6'	d	5.8	3.8	4.6	5.8	5.8	3.8
	7pR	i	2.1	4.8	4.2	3.4	4.2	6.5
	7pS	NO	3.8	5.8	2.7	4.9	2.7	5.0
	Gal2	m	3.3	9.2	6.2	4.1	5.00	9.9
Fuc6	7pR	d	3.4	7.1	6.7	4.0	4.5	8.6
Gal(1→3)GlcNAc			<i>syn-ψ</i> (+)	<i>syn-ψ</i> (+)	<i>anti-ψ</i>	<i>syn-ψ</i> (+)	<i>anti-ψ</i>	<i>anti-ψ</i>
Fuc(1→4)GlcNAc			<i>syn-ψ'</i> (+)	<i>anti-ψ'</i>	<i>syn-ψ'</i> (+)	<i>syn-ψ'</i> (-)	<i>syn-ψ'</i> (-)	<i>anti-ψ'</i>

Tabla 4.10: Distancias interprotónicas (Å) para cada uno de los conformeros A, D, E, B, H e I junto con la intensidad NOEs observados. Por simplicidad, en esta tabla no se incluyen los mínimos F y G, ya que los datos de constantes de acoplamiento han permitido descartar que estos conformeros estuvieran significativamente poblados en disolución. Las distancias Fuc1-Glc6NAc, Fuc1-Glc6NAc', Fuc5-Glc6NAc y Fuc5-Glc6NAc' se dedujeron a partir del rotámero más estable (gt).

Por otro lado, la ausencia del NOE Gal1-GlcNAc4, exclusivo de la familia conformacional *anti- ψ* , representada por los mínimos **E**, **H** e **I**, indicaría que estas conformaciones apenas están pobladas en disolución. Este resultado es consistente con los resultados de mecánica molecular para los conformeros **H** e **I** pero está en desacuerdo para el conformero **E**, para el que se predijo una población del 14%.

En cuanto a la familia conformacional *syn- ψ'* (-), representada por el mínimo **B**, y que fue previamente detectada en el Le^a, el hecho de que los NOE Fuc1-H7_{PR} y NOE Fuc5-GlcNAc3 sean de intensidad débil indica que, si este conformero estuviera presente en disolución, su población debería ser muy pequeña, de acuerdo con los cálculos de mecánica molecular. Por tanto, simplemente con la presencia de los conformeros **A** y **D** en disolución, sería posible justificar los NOEs observados.

De igual manera que en el caso del Le^a, se procedió a estimar la población de cada uno de ellos, obteniéndose el valor más pequeño del factor R usando una proporción **A:D** de 90:10, como se representa en la **Figura 4.13**.

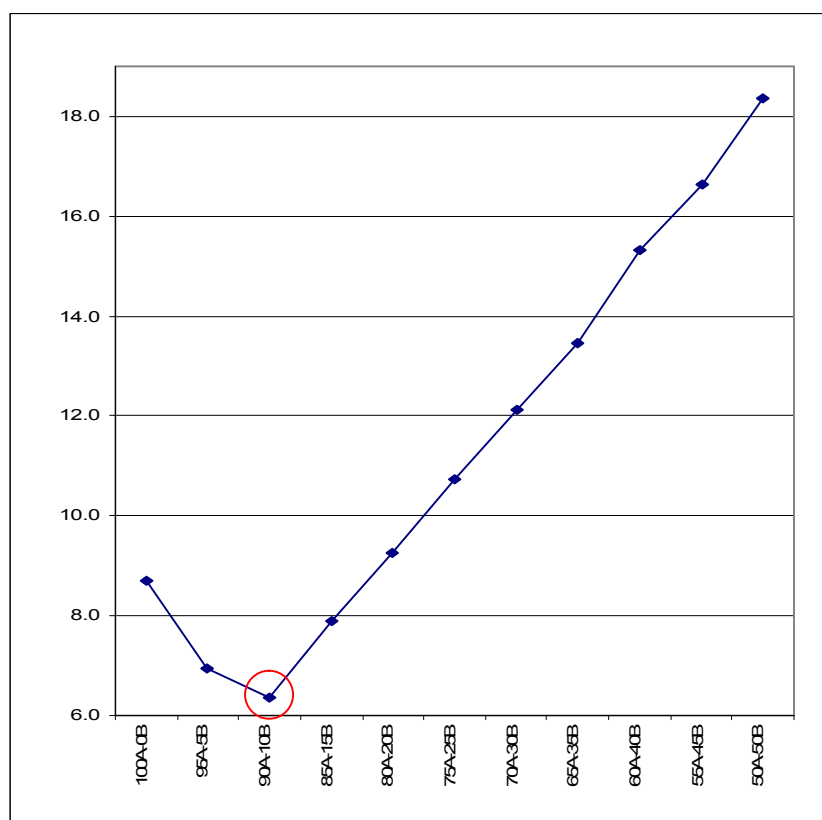


Figura 4.13: Representación del error en los NOEs en función de las distintas proporciones de conformeros.

También se puede realizar una estimación de la población de cada conformero teniendo en cuenta los datos experimentales de constantes de acoplamiento. En este caso se aplicó la misma metodología se será explicada en el capítulo 5, obteniendo un mejor ajuste entre datos teóricos y experimentales para una proporción 86:14 de **A:D**. Este resultado cuadra bastante bien con el obtenido a partir de los NOEs (**Figura 4.14**).

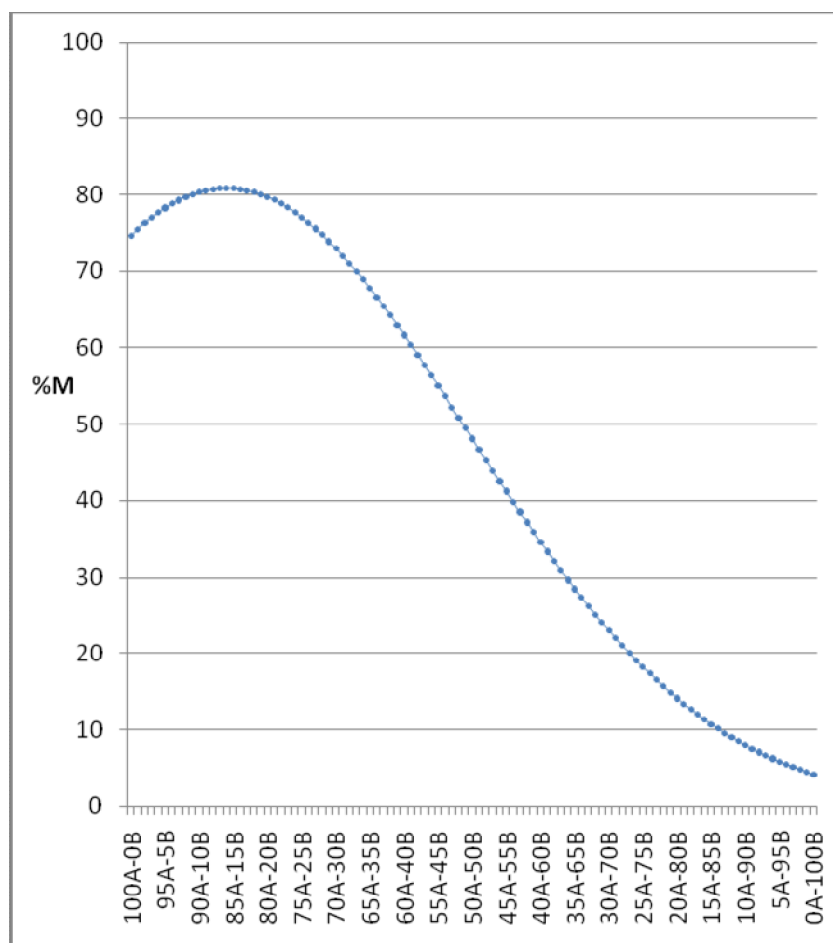


Figura 4.14: Representación del grado de ajuste (%M) entre valores teóricos y experimentales frente a la población de distintas proporciones de conformeros **A** y **D**.

3.3. Conclusiones:

El análisis conformacional del C-Le^a en disolución acuosa mediante RMN y cálculos de mecánica molecular indica la existencia de dos familias conformacionales en disolución, que se diferencian en la conformación en torno al enlace aglicónico entre las unidades de fucosa y la N-acetilglucosamina. Este enlace adopta una disposición *syn*-periplanar en el conformero mayoritario (*syn-ψ(+)-synψ'* (+)), y *anti*-periplanar en el conformero minoritario (*syn-ψ(+)-anti-ψ'*), lo que supone una diferencia conformacional con respecto al trisacárido Le^a, en el que solamente está poblada la región *syn* (*synψ(+)* o *synψ'(-)*), en proporción 95:5).

Mientras que en los casos anteriores, en los que se había comparado el comportamiento conformacional de los O- y C-glicósidos, se había visto que la sustitución del oxígeno puente por un grupo CH₂ producía cambios conformacionales en la unión interglicosídica modificada, en este caso, la sustitución del oxígeno puente entre las unidades de galactosa y N-acetilglucosamina prácticamente no altera el comportamiento conformacional en torno a esta unión, pero por el contrario, sí provoca un cambio conformacional en torno a la unión interglicosídica contigua entre las unidades de fucosa y N-acetilglucosamina con la presencia de la familia conformacional *anti-ψ'* en disolución.

La comparación del Le^a y el C-Le^a sirve para arrojar luz sobre los motivos de la rigidez conformacional del determinante Le^a, que ha sido atribuida a una combinación de factores electrónicos y estéricos. Mientras que en torno al enlace Gal→GlcNAc, en el que desaparece el efecto *exo*-anomérico no se observan cambios conformacionales apreciables, si que se detectan alrededor del enlace Fuc→GlcNAc en el que el efecto *exo*-anomérico se mantiene. Es decir, los resultados indican que el comportamiento de ambos enlaces está correlacionado y que los factores estéricos son decisivos para modular la geometría y dinámica de esta molécula.

4. Parte Experimental:

4.1. Materiales:

El compuesto **Lewis A** está disponible comercialmente como mezcla anomérica (**3**) y bloqueado en su forma β (**3'**) en las compañías Dextra y Toronto, respectivamente. El compuesto **4**, obtenido por síntesis química, fue suministrado por el grupo de P. Sinnay (París, Francia).

4.2. Mecánica Molecular:

El espacio conformacional del Le^a y C-Le^a se exploró mediante una búsqueda conformacional estocástica utilizando la opción “*Mixed torsional/Low mode Sampling*” disponible en Maestro v8.0, eligiendo MM3* como campo de fuerzas y una constante dieléctrica de 80 debyes para simular el medio acuoso. Se seleccionaron las torsiones glicosídicas y aglicónicas, como “torsiones de rotación” y los esqueletos carbonados de los anillos como átomos comparativos, fijando un valor de 0.50 Å de RMSD para considerar dos conformeros como iguales. Se realizaron 20000 pasos y se estableció una ventana de energía de 21 kJ/mol para seleccionar los conformeros. Cada una de las estructuras generadas fue minimizada mediante 500 iteraciones, usando gradientes conjugados (TNGC), hasta llegar a la convergencia.

Tras los cálculos, se obtuvieron 383 estructuras para **3** y 949 estructuras para **4**. Para identificar el número de familias conformacionales se utilizó el módulo X-Cluster, incluido también en MAESTRO v8.0. A través de este protocolo, fue posible definir los conformeros relevantes de **3** mediante tres geometrías. En el caso de **4**, se obtuvieron 8 familias conformacionales diferentes.

4.3. Cálculos de constantes de acoplamiento y NOEs teóricos:

Las constantes de acoplamiento fueron calculadas para cada conformación del compuesto **4** usando la ecuación de Karplus-Altona¹⁸. Los valores de la distribución se calcularon de acuerdo con la ecuación 2.1:

$$J = \sum P_{\phi,\psi} \times J_{i\phi,\psi} \quad \text{Ecuación 2.1}$$

Las distancias interprotónicas promedio se calcularon aplicando la siguiente expresión:

$$\langle r^{-6} \rangle_{kl} = \sum P_{\phi, \psi} \times r^{-6}_{kl(\phi, \psi)} \quad \text{Ecuación 2.2}$$

Las intensidades teóricas de los NOEs se determinaron de acuerdo con la matriz de relajación completa, empleando el programa NOEPROM. Se asumió un movimiento isotrópico de las moléculas y un valor de relajación externa de 0.1 s^{-1} . Se fijaron los valores de tiempo de correlación en 38 y 34 ps, para **3** y **4**, respectivamente, por ser los valores que mejor ajustaban los valores teóricos y experimentales de los NOEs intra-residuos Gal1-Gal3, Fuc1-Fuc2 y Fuc1-Fuc5, que no dependen del valor de los ángulos glicosídicos.

4.4. Experimentos de RMN:

Los espectros del compuesto **3**, **3'** y **4** se adquirieron en un instrumento Bruker Avance 500 MicroBay, equipado con una criosonda de 5mm. Se usaron concentraciones 7 y 4 mM para los compuestos comerciales **3** y **3'**, respectivamente. Los experimentos del compuesto **4** se realizaron empleando una concentración 3mM. Los experimentos se llevaron a cabo a 298 K, en D_2O , utilizando DSS como referencia externa.

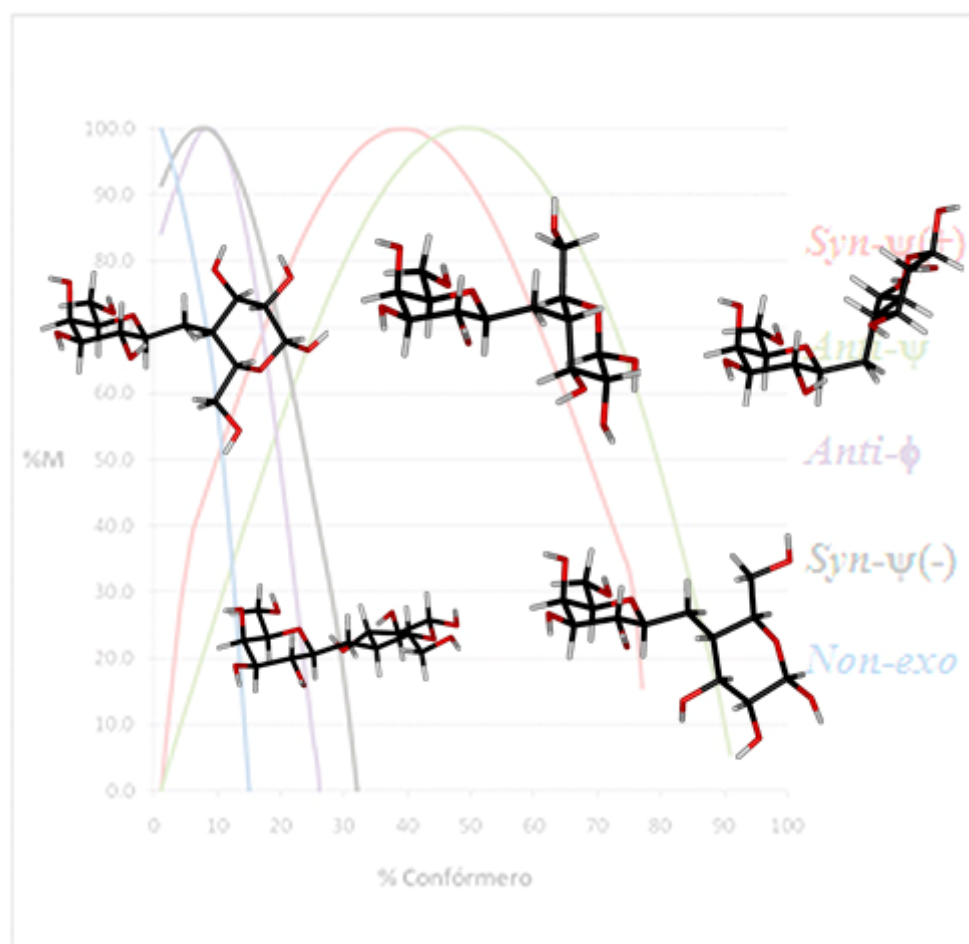
Para la asignación de los compuestos **3**, **3'** y **4** se registraron experimentos de ^1H , COSY¹⁹, HSQC²⁰ y 1D-TOCSY²¹ (80 ms de tiempo de mezcla). Para la determinación de la estructura tridimensional, se emplearon experimentos 1D y 2D-NOESY¹⁵ y 2D-ROESY²². En todos los experimentos 2D, se usaron tiempos de mezcla de 500 y 300 ms. mientras que, para los experimentos 1D, se seleccionaron tiempos de mezcla de 500 ms.

5. Bibliografía

- (1) Küber., L. *Wien Klin Wochenschr* **1901**, *14*, 1132.
- (2) Lloyd, K. O. *Glycoconj J* **2000**, *17*, 531.
- (3) Feizi, T. *Trends Biochem Sci* **1991**, *16*, 84.
- (4) Lemieux, R. U.; Bock, K.; Delbaere, L. T. J.; Koto, S.; Rao, V. S. *Can J Chem* **1980**, *58*, 631.
- (5) Cagas, P.; Bush, C. A. *Biopolymers* **1990**, *30*, 1123.
- (6) Mukhopadhyay, C.; Bush, C. A. *Biopolymers* **1991**, *31*, 1737.
- (7) Kogelberg, H.; Rutherford, T. J. *Glycobiology* **1994**, *4*, 49.
- (8) Imberty, A.; Mikros, E.; Koca, J.; Mollicone, R.; Oriol, R.; Perez, S. *Glycoconj J* **1995**, *12*, 331.
- (9) Csonka, G. I.; Sosa, C. P.; Csizmadia, I. G. *J. Phys. Chem. A* **2000**, *104*, 3381.
- (10) Azurmendi, H. F.; Martin-Pastor, M.; Bush, C. A. *Biopolymers* **2002**, *63*, 89.
- (11) Bekiroglu, S.; Kenne, L.; Sandström, C. *Carbohydr. Res.* **2004**, *339*, 2465.
- (12) Geyer, A.; Hummel, G.; Eisele, T.; Reinhardt, S.; Schmidt, R. *Chem Eur J* **1996**, *2*, 981.
- (13) Geyer, A.; Reinhardt, S.; Bengas, G.; Rothe, U.; Schmidt, R. *J Am Chem Soc* **1997**, *119*, 11708.
- (14) Berthault, P.; Birlirakis, N.; Rubinstenn, G.; Sinaÿ, P.; Desvaux, H. *J Biomol NMR* **1996**, *8*, 23.
- (15) Neuhaus, D. W., M. P. *The Nuclear Overhauser Effect in Structural and Conformational Analysis*; VCH: New York, 1989.
- (16) Gonzalez, C.; Rullmann, J. A. C.; Bonvin, A. M. J. J.; Boelens, R.; Kaptein, R. *J. Magn. Reson. (1969)* **1991**, *91*, 659.
- (17) Poveda, A.; Asensio, J. L.; Martín-Pastor, M.; Jiménez-Barbero, J. *J Biomol NMR* **1997**, *10*, 29.
- (18) Altona, C.; Sundaral.M *J Am Chem Soc* **1972**, *94*, 8205.
- (19) Aue, W. P.; Bartholdi, E.; Ernst, R. R. *J Chem Phys* **1976**, *64*, 2229.
- (20) Willker, W.; Leibfritz, D.; Kerssebaum, R.; Bermel, W. *Magn Reson Chem* **1993**, *31*, 287.
- (21) Adell, P.; Parella, T.; Sanchezferrando, F.; Virgili, A. *J. Magn. Reson. Series B* **1995**, *108*, 77.
- (22) Bothnerby, A. A.; Stephens, R. L.; Lee, J. M.; Warren, C. D.; Jeanloz, R. W. *J Am Chem Soc* **1984**, *106*, 811.

Capítulo 5

JAMFIS: Una nueva herramienta para el estudio conformacional de C-disacáridos.



1. Introducción y antecedentes:

Como se ha explicado en el capítulo 2 de esta memoria, a finales del siglo pasado existió un vivo debate sobre el grado de similitud conformacional entre los oligosacáridos naturales y sus análogos *C*-glicosídicos. Los estudios iniciales de Kishi en la Universidad de Harvard¹⁻¹⁶ basados en constantes de acoplamiento protón-protón, habían propuesto que el comportamiento conformacional de los *C*-disacáridos era comparable al de los correspondientes *O*-disacáridos. Por otra parte, estudios más exhaustivos de algunos *C*-disacáridos mediante la combinación de cálculos de mecánica molecular y datos de RMN (constantes de acoplamiento y NOEs) realizados por nuestro grupo de investigación¹⁷ pusieron de manifiesto, en muchos de los casos estudiados, la existencia de diferencias conformacionales significativas entre ambos tipos de compuestos. Uno de estos casos es el del *C*-disacárido $\beta(1\rightarrow3)$ Gal-Glc, cuyo análisis conformacional se ha descrito en el capítulo 2 de esta tesis. Por tanto, nuestro grupo de trabajo demostró que las conclusiones de Kishi y sus colaboradores no tenían validez general y que era necesario estudiar cada caso independientemente.

La aproximación seguida por Kishi para el análisis conformacional de los *C*-disacáridos consistía en situar al compuesto en una red de diamante, para así visualizar las interacciones estéricas desfavorables que tienen lugar en cada una de las conformaciones alternadas posibles y, a partir de las constantes de acoplamiento protón-protón, deducir la conformación mayoritaria del *C*-disacárido. Estos estudios llevaron al grupo de Kishi a concluir que los *C*-disacáridos presentaban una gran preferencia por la conformación *exo/syn* en torno al enlace pseudo-glicosídico para evitar interacciones *gauche* desestabilizantes, mientras en torno al enlace pseudo-aglicónico adoptaban conformaciones alternadas, a ser posible, libres de interacciones 1,3-diaxiales desestabilizantes.

En uno de sus trabajos, Kishi argumentó que era difícil realizar el análisis conformacional de *C*-disacáridos mediante datos de NOE ya que, debido a la dependencia del NOE con la inversa de la sexta potencia de la distancia interprotónica, pequeños errores en la cuantificación de la intensidad del NOE podían conducir a errores significativos en la caracterización del equilibrio

conformacional¹⁰. Hoy en día, es obvio que este punto de vista es extremo, ya que la combinación de NOEs y constantes de acoplamiento ha permitido el estudio detallado de las propiedades conformacionales de un gran número de oligosacáridos, de moléculas orgánicas, y de otras biomoléculas en disolución¹⁸. No obstante, es indudable que la correspondencia directa entre intensidad de NOE y distancias promedio no es evidente, ya que el NOE también depende de la movilidad molecular y del entorno químico de la pareja de protones en cuestión. Además, el promediado de los NOES en equilibrios conformacionales no es lineal con el peso relativo de las conformaciones. Aunque la relación entre constantes de acoplamiento escalares y ángulos de torsión también es compleja, ya que las Js dependen de la electronegatividad de los sustituyentes a lo largo del camino de acoplamiento y de la orientación relativa de los mismos, en equilibrios conformacionales, su promediado sí es lineal con el peso relativo de las conformaciones en equilibrio.

Muchas moléculas importantes presentan equilibrios conformacionales en torno a distintos grados de libertad de torsión. En este sentido, en este capítulo hemos decidido analizar el comportamiento conformacional de distintos C-disacáridos utilizando una aproximación cuantitativa basada exclusivamente en datos experimentales de constantes de acoplamiento escalares. Nuestro análisis está basado en el método propuesto por D. Cicero y colaboradores, denominado NAMFIS (*NMR Analysis of Molecular Flexibility In Solution*)¹⁹ para el análisis conformacional de moléculas flexibles en disolución, que consiste en determinar la distribución de población de un conjunto de conformeros dado que presenta un mayor acuerdo con los datos de NOE y constantes de acoplamiento. En nuestro caso, nuestra hipótesis se basa en que, de hecho, el comportamiento conformacional de los C-disacáridos está contenido en los valores de las cuatro constantes de acoplamiento experimentales protón-protón medibles alrededor de los dos enlaces pseudo-glicosídicos. Por tanto, un análisis riguroso de estos valores, apoyado por modelos geométricos fidedignos, proporcionados por métodos de modelado molecular, debería permitir la descripción cuantitativa del equilibrio conformacional alrededor de ambos ángulos de torsión. Al método lo hemos denominado JAMFIS (*J-based Analysis of Molecular Flexibility in Solution*) y puede ser de validez general para describir equilibrios conformacionales en los casos en que existan suficientes datos experimentales de Js que no estén correlacionados entre sí.

2. Método JAMFIS:

El proceso seguido puede dividirse en cinco etapas:

1. Se realiza un estudio conformaciones de la molécula objetivo, mediante cálculos de mecánica molecular, para determinar el conjunto de conformeros energéticamente posibles que podrían estar presentes en la distribución final. Una vez seleccionados los conformeros, sus energías relativas no se tienen en cuenta en el resto del proceso, por lo que las poblaciones finales no dependerán de la energía de cada conformero que viene dado por la ecuación del campo de fuerzas utilizado.
2. Se determinan las constantes de acoplamiento teóricas de cada conformero mediante la ecuación de Karplus-Altona²⁰.
3. Se generan todas las distribuciones de población posibles por combinación de los diferentes conformeros seleccionados y se calculan las constantes de acoplamiento promedio que se obtendrían en cada caso. Para ello se tienen en cuenta las siguientes reglas:
 - La suma de todas las poblaciones no puede exceder del 100%.
 - Para reducir el número de combinaciones posibles, sólo se tienen en cuenta número enteros para las poblaciones de cada conformero.
 - Sólo se consideran válidas aquellas soluciones en las que los valores calculados de las constantes de acoplamiento promedio estén comprendidos en un rango determinado:

$$J_i^{\text{EXP}} - \Delta J_i \leq J_i^{\text{CAL}} \leq J_i^{\text{EXP}} + \Delta J_i$$

Donde ΔJ_i corresponde al error asociado a la comparación entre constantes de acoplamiento experimentales y teóricas. Hemos estimado que este error es de ± 1.5 Hz, considerando tanto el error en la medida experimental (~ 0.5 Hz) como el error en la ecuación de Karplus-Altona (~ 1.0 Hz).

4. Se examina el grado de ajuste entre valores experimentales y teóricos para cada distribución de población mediante el empleo de un parámetro, M que viene

dado por la siguiente expresión matemática, y que puede adoptar valores entre cero y uno.

$$M = \prod_i e^{-\left\{ \frac{1}{2} \left[\frac{J_i^{\text{exp}} - J_i^{\text{Cal}}}{\Delta J_i} \right]^2 \right\}}$$

5. Para la interpretación de los datos se siguen tres aproximaciones diferentes:

- Método 1. Se determina la distribución de población que proporciona el mejor ajuste entre datos teóricos y experimentales (mayor valor de M).
- Método 2. Se representa la población de cada conformero frente al valor de M normalizado (el máximo valor de M corresponde al 100% y el menor al 0%). Para cada población, se elige el mejor valor de M de un conformero dado, independientemente de la población del resto de conformeros. Esta representación da una idea de todas las soluciones que son consistentes con las constantes de acoplamiento, y permite evitar la sobre-interpretación de datos que podría producirse si sólo se tuviera en cuenta una solución única. A partir de estas gráficas es posible visualizar como varía el ajuste entre datos teóricos y experimentales con la población de cada conformero. Las curvas obtenidas pueden dividirse en tres tipos que se representan en la **Figura 5.1**.

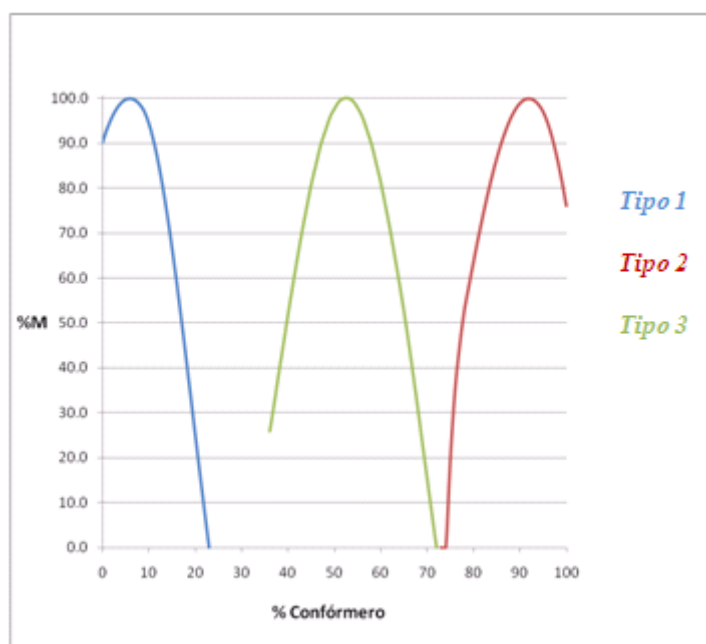


Figura 5.1: Tipo de curvas obtenidas en la representación de la población de cada conformero frente al valor de M.

a) Las curvas de tipo 1 (en azul en la **Figura 5.1**) son aquellas que presentan un valor del parámetro M del 100% o cercano al 100% cuando el confórmero no está poblado. En estos casos, la presencia del confórmero en disolución no puede demostrarse por constantes de acoplamiento.

b) Las de tipo 2 (en rojo) son aquellas que presentan un valor del parámetro M del 100% o cercano al 100%, cuando la población del confórmero es del 100%. En estos casos los valores de las constantes de acoplamiento prácticamente se podrían justificar con una única familia conformacional.

c) Las de tipo 3 (en verde) son las intermedias entre ambas situaciones, que corresponden a confórmeros que deben estar poblados en disolución para que se pueda obtener un buen ajuste, pero que requieren la presencia de al menos otro confórmero para poder explicar los valores experimentales de las constantes de acoplamiento

Además, la anchura de las curvas proporciona una idea de lo bien definida que está la distribución de población deducida a partir únicamente de constantes de acoplamiento protón-protón. En aquellos casos en que las curvas fueran anchas, sería necesario recurrir a otros datos experimentales, por ejemplo NOEs, para determinar las poblaciones de cada confórmero con mayor exactitud, mientras que si fuesen estrechas, estarían indicando que las constantes de acoplamiento, por sí solas, serían suficientes para caracterizar el comportamiento conformacional del C-disacárido en cuestión.

- Método 3. Se representa el valor del parámetro M frente al número de confórmeros que se tiene en cuenta en su cálculo. En cada caso, se seleccionan los confórmeros que proporcionan el mayor valor de M. De esta forma, se puede determinar el número mínimo de familias conformacionales que deben estar pobladas para conseguir un buen ajuste.

En este capítulo, hemos seleccionado un conjunto amplio de C-disacáridos para los que existen datos experimentales de constantes de acoplamiento. En primer lugar, hemos aplicado el método JAMFIS a dos C-disacáridos estudiados previamente por nuestro grupo de investigación (C-lactosa²¹⁻²² y α -C-manobiosa²³) y al disacárido

estudiado en el capítulo 2 de esta tesis (β -C-Gal-(1 \rightarrow 3)- β -Glc-OMe). Así, hemos comparado los resultados de JAMFIS con los datos obtenidos con una combinación "clásica" de mecánica molecular, NOEs y constantes de acoplamiento. A continuación, hemos analizado otros C-disacáridos estudiados previamente por Kishi y sus colaboradores. Hemos dividido estas moléculas según el tipo de unión pseudo-glicosídica que presentan (1 \rightarrow 1), (1 \rightarrow 3) y (1 \rightarrow 4). A partir de aquí, al referirnos al enlace pseudo-glicosídico y pseudo-aglicónico de un C-disacárido omitiremos el prefijo *pseudo* y utilizaremos, por simplicidad, los términos glicosídico y aglicónico correspondientes a los O-disacáridos naturales.

3. Aplicación a C-disacáridos estudiados mediante mecánica molecular, NOEs y constantes de acoplamiento.

3.1 C-Lactosa:

La C-lactosa (β -C-Gal(1 \rightarrow 4)-Glc) es el C-disacárido que ha sido estudiado más en detalle, al menos por tres grupos de investigación diferentes, por lo que se cuenta con multitud de datos experimentales. Como se comentó en el capítulo 2, nuestro grupo de investigación^{21,23} y el grupo de Berthault²⁴, de manera independiente, analizaron el comportamiento conformacional de la C-lactosa, llegando a la conclusión de que los datos de RMN sólo se podían explicar con la existencia de dos familias conformacionales mayoritarias alrededor de enlace aglicónico (*syn- ψ* y *anti- ψ*). Además, nuestro grupo de investigación detectó la presencia de una pequeña población del confórmero *anti- ϕ* . Estos resultados pusieron de manifiesto la mayor flexibilidad del C-disacárido con respecto a la lactosa donde el mínimo *syn- ψ* es muy predominante.

Para la aplicación del método JAMFIS se utilizaron como punto de partida los cinco mínimos obtenidos mediante cálculos de mecánica molecular (*syn- ψ (+)*, *anti- ψ* , *anti- ϕ* , *syn- ψ (-)* y *non exo*, **Figura 5.2**) y las constantes de acoplamiento inter-glicosídicas medidas en nuestro grupo de investigación²¹ que son similares a las obtenidas por el grupo de Kishi¹⁶ (**Tabla 5.1**).

	Jiménez-Barbero <i>et al.</i>	Kishi <i>et al.</i>
Gal1'-H7pR	10.3	9.9
Gal1'-H7pS	1.2	~0.2
Glc4-H7pR	3.5	3.4
Glc4-H7pS	5.2	4.7

Tabla 5.1: Constantes de acoplamiento inter-glicosídicas medidas para la C-lactosa.

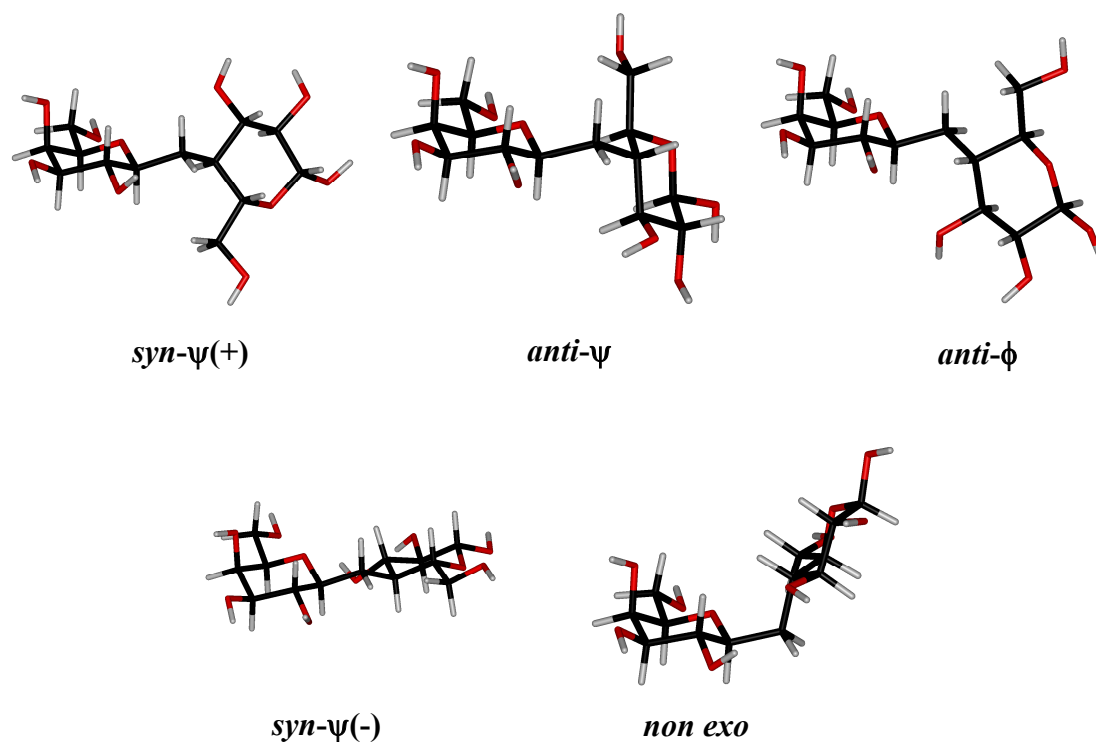


Figura 5.2: Conórmeros de la C-lactosa.

La aplicación del método JAMFIS reveló que existen 280694 distribuciones de población para las que se obtienen constantes de acoplamiento promedio que son compatibles con las constantes experimentales, según las reglas establecidas en la sección anterior (hasta un error de ± 1.5 Hz en cada constante). De todas estas soluciones, la distribución de población *anti-ψ:syn-ψ(+):syn-ψ(-):anti-φ:non exo* que proporciona el mejor ajuste es 48:38:7:7:0 con un factor M de 0.972, que es muy similar a la que se predice con el campo de fuerzas MM3* (**Tabla 5.2**) y que está de acuerdo con los NOEs observados.

	MM3*	JAMFIS
<i>anti-ψ</i>	56	48
<i>syn-ψ(+)</i>	30	38
<i>syn-ψ(-)</i>	7	7
<i>anti-φ</i>	3	7
<i>non exo</i>	4	0

Tabla 5.2: Resultados de la distribución calculada con el campo de fuerzas MM3* y la obtenida con el método JAMFIS para la C-lactosa.

En el diagrama que se representa en la **Figura 5.3** se comparan los valores de las constantes de acoplamiento de la mejor solución de JAMFIS con los calculados para la distribución teórica MM3* y los medidos experimentalmente.

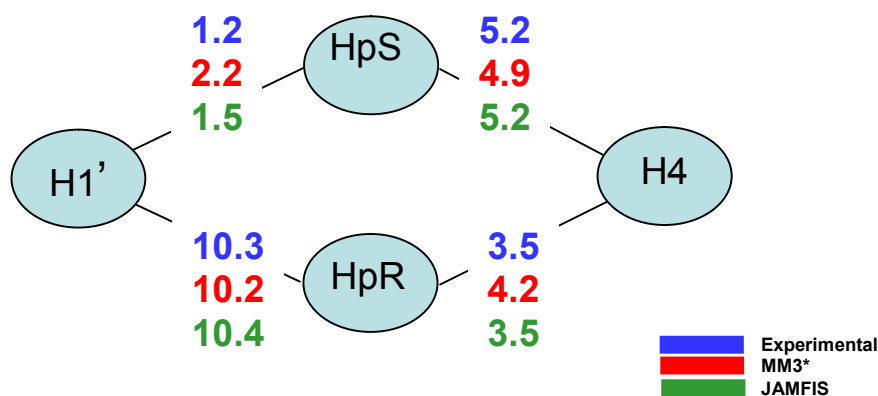


Figura 5.3: Diagrama representativo de las constantes de acoplamiento medidas experimentalmente para la C-lactosa, junto a las obtenidas mediante cálculos de mecánica molecular (MM3*) y mediante el método JAMFIS.

En el segundo método de análisis, se representó la población de cada conformero frente a M normalizado, tal como se describe en la sección anterior. (**Figura 5.4**)

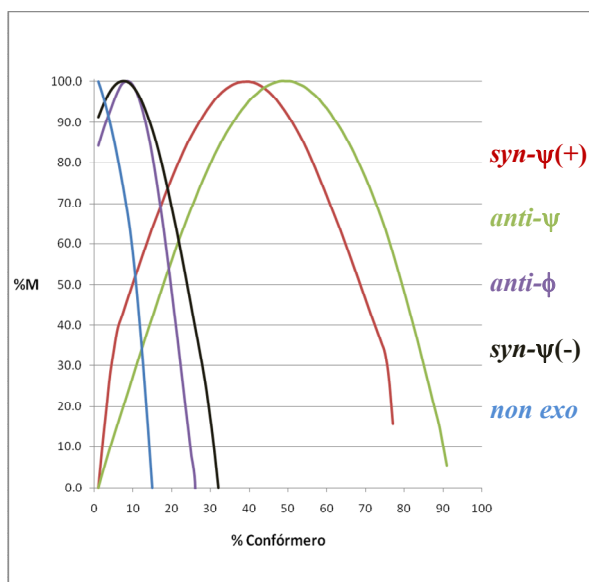


Figura 5.4: Representación del valor de M normalizado frente a la población de cada conformero de la C-lactosa.

De la inspección de las gráficas se puede concluir que hay dos familias conformacionales *syn-ψ(+)* y *anti-ψ*, que están significativamente pobladas en

disolución (curvas de tipo 3), frente a la predominancia del conformero *syn* que caracteriza al disacárido natural. El resto de conformeros presentan curvas de tipo 1, por lo que es posible conseguir buenos ajustes sin estos conformeros, por lo que, o no están presentes en disolución o su población es pequeña.

Si se analiza el valor de M en función del número de conformeros tenidos en cuenta en el cálculo (Método 3), obtenemos la gráfica de la **Figura 5.5** que nos permite ver la mejora del ajuste al aumentar el número de conformeros. Si se considera un sólo conformero (*anti- ψ*) se obtiene un valor de M muy bajo (0.37), por lo que con una sólo familia conformacional no es posible obtener un buen ajuste. El círculo representado en la gráfica indica que las constantes calculadas no son compatibles con los valores experimentales (error mayor de 1.5 Hz en al menos una de las constantes). El valor de M aumenta a 0.78 al introducir una población del 36% del conformero *syn- ψ (+)*, poniéndose de manifiesto que para conseguir un ajuste razonable se necesita al menos la presencia de dos familias conformacionales en disolución. Además, si se considera una población del 8% de un tercer conformero (*anti- ϕ*) el ajuste mejora, con un valor de M por encima de 0.9. Esta distribución de población de tres conformeros (*anti- ψ :syn- ψ (+):anti- ϕ* 58:34:8) explica perfectamente los NOEs observados en este C-disacárido, en concreto, NOEs 1'→4 y 1'→3 intensos (característicos de los conformeros mayoritarios *syn- ψ* y *anti- ψ* respectivamente) y un NOE 2'→4 poco intenso, que al ser exclusivo del conformero *anti- ϕ* indicaría la presencia de una pequeña población de esta familia conformacional.

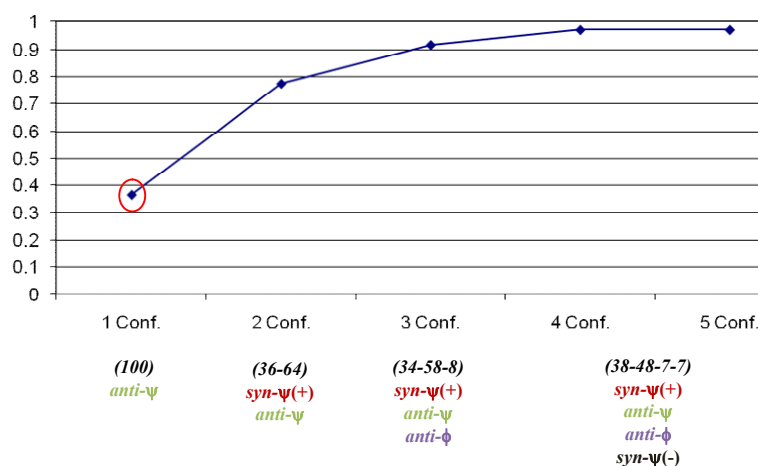


Figura 5.5: Representación del factor M en función del número de conformeros utilizados en el método JAMFIS aplicado a la C-lactosa. En cada caso se seleccionan los conformeros que dan un mayor valor de M y las poblaciones se indican entre paréntesis.

3.2 α -C-Manobiosa:

Como en la C-lactosa, el punto de partida del análisis del α -C-manobiósido (α -C-Man(1 \rightarrow 2)- α -Man-OMe) son los mínimos energéticos obtenidos con el campo de fuerzas MM3* y las constantes de acoplamiento inter-glicosídicas medidas por nuestro grupo de investigación²³ (Tabla 5.3). En este caso los cálculos indicaron la existencia de tres conformeros posibles *syn*- ψ (-), *syn*- ψ (+) y *non exo* (Figura 5.6).

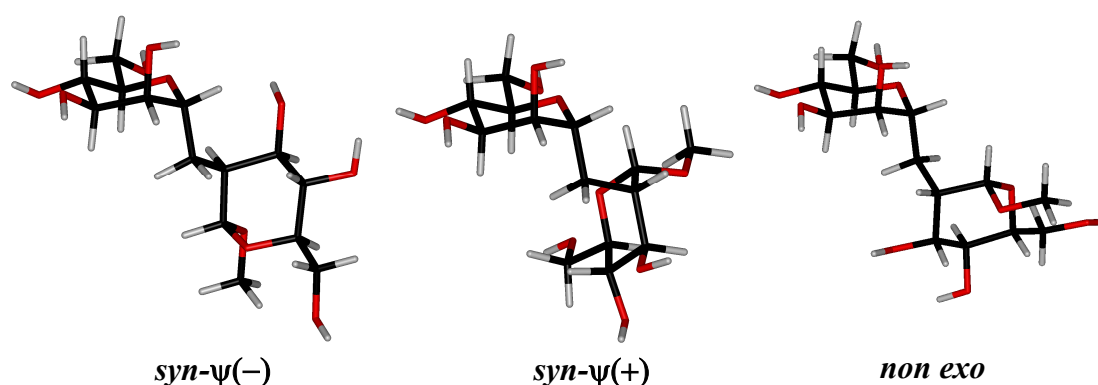


Figura 5.6: Conformeros posibles para el α -O-metil- α -C-manobiósido.

	Espinosa <i>et al.</i>
Man1'-H7pS	8.7
Man1'-H7pR	3.5
Man2-H7pS	8.9
Man2-H7pR	4.1

Tabla 5.3: Constantes de acoplamiento inter-glicosídicas del α -O-metil- α -C-manobiósido.

	MM3*	JAMFIS
<i>syn</i> - ψ (-)	16	25
<i>syn</i> - ψ (+)	59	52
<i>non exo</i>	25	23

Tabla 5.4: Poblaciones calculadas para el α -O-metil- α -C-manobiósido con el campo de fuerzas MM3* y con el método JAMFIS.

Siguiendo el mismo protocolo que en el caso anterior, se encontraron 331 soluciones que son compatibles con las constantes de acoplamiento experimentales. De todas estas soluciones, la distribución de población *syn*- ψ (-):*syn*- ψ (+):*non exo* que mejor se ajusta a los datos experimentales con un factor M de 0.694 es 25:52:23, en excelente acuerdo con la distribución calculada con el campo de fuerzas MM3* (Tabla 5.4)

En el siguiente diagrama se han representado las constantes de acoplamiento de la mejor solución de JAMFIS junto a los valores de la distribución calculada con MM3* y los medidos experimentalmente (**Figura 5.7**).

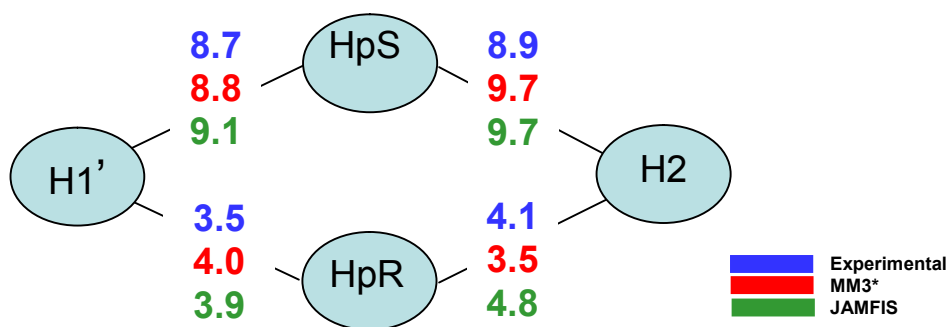


Figura 5.7: Diagrama representativo de las constantes de acoplamiento del α -O-metil- α -C-manobiósido.

A continuación se representó la población de cada conformero frente al valor de M para visualizar todas las soluciones posibles (**Figura 5.8**). Como puede observarse las curvas de los tres conformeros son de tipo 3, por lo que para explicar las constantes de acoplamiento experimentales es necesario que las tres familias conformacionales se encuentren pobladas en disolución, descartándose la existencia de una conformación única. Además las curvas son más estrechas que en el caso de la C-lactosa, por lo que el equilibrio conformacional está mejor definido.

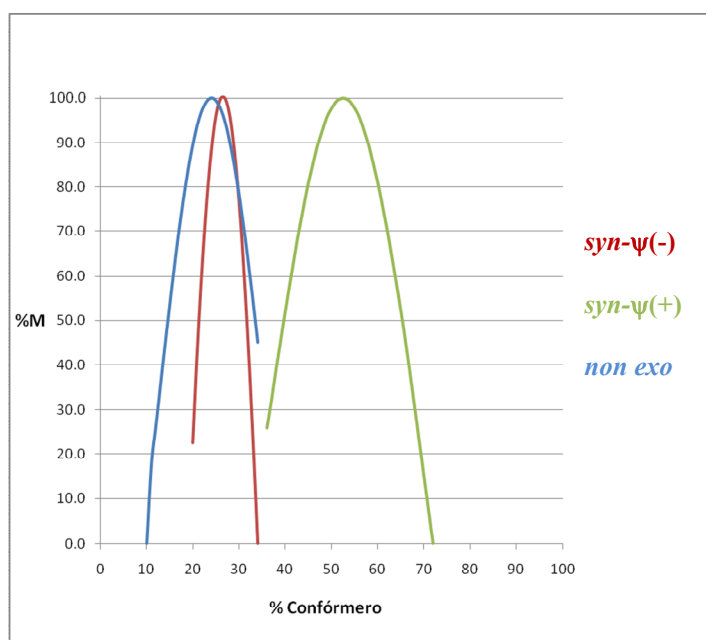


Figura 5.8: Representación del valor de M normalizado frente a la población de cada conformero del α -O-metil- α -C-manobiósido.

Al igual que en el caso anterior, si se analiza el valor de M en función del número de confórmeros, se obtiene la gráfica de la **Figura 5.9**. Como puede observarse sólo se consiguen valores aceptables de M cuando se tienen en cuenta tres confórmeros, confirmándose de nuevo que las tres familias conformacionales están presentes en disolución. Como en el caso anterior el círculo en la gráfica representa que ni con uno ni con dos confórmeros se obtienen constantes de acoplamiento compatibles con los valores experimentales.

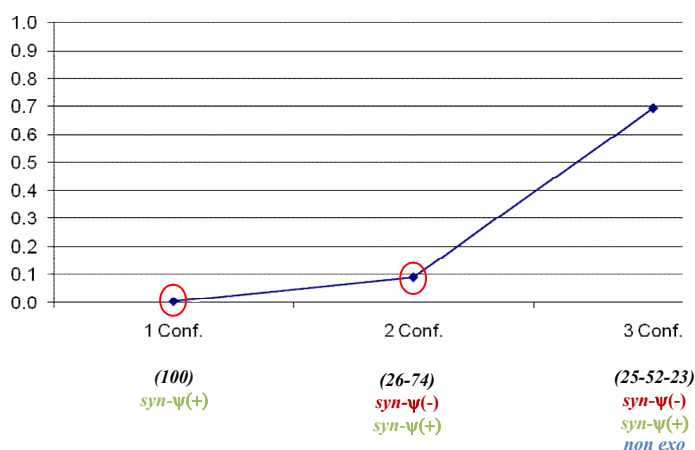


Figura 5.9: Representación de la población de cada confórmero frente al valor de M para el α -O-metil- α -C-manobiósido.

Los datos de NOE que se obtuvieron en el estudio de RMN previo²³ (NOEs intensos Man1'-Man2 y Man5'-Man1 característicos del confórmero mayoritario $syn-\psi(+)$ y NOEs Man5'-Man2 y Man2'-Man2 exclusivos de los confórmeros minoritarios $syn-\psi(-)$ y $non\ exo$ respectivamente) son consistentes con la distribución de población obtenida con JAMFIS. Por tanto, los NOEs sirven simplemente para confirmar las conclusiones obtenidas mediante constantes de acoplamiento pero no añaden información complementaria.

3.3 β -C-Gal-(1 \rightarrow 3)- β -Glc-OMe:

El tercer disacárido al que se le aplicó esta aproximación fue el estudiado previamente en el capítulo 2 de esta Tesis (β -C-Gal-(1 \rightarrow 3)- β -Glc-OMe). Como ya vimos, los cálculos predijeron la existencia de 5 familias conformacionales distintas que se utilizaron como punto de partida para el método JAMFIS (**Figura 5.10**) junto a las constantes de acoplamiento medidas para este C-disacárido (**Tabla 5.5**)

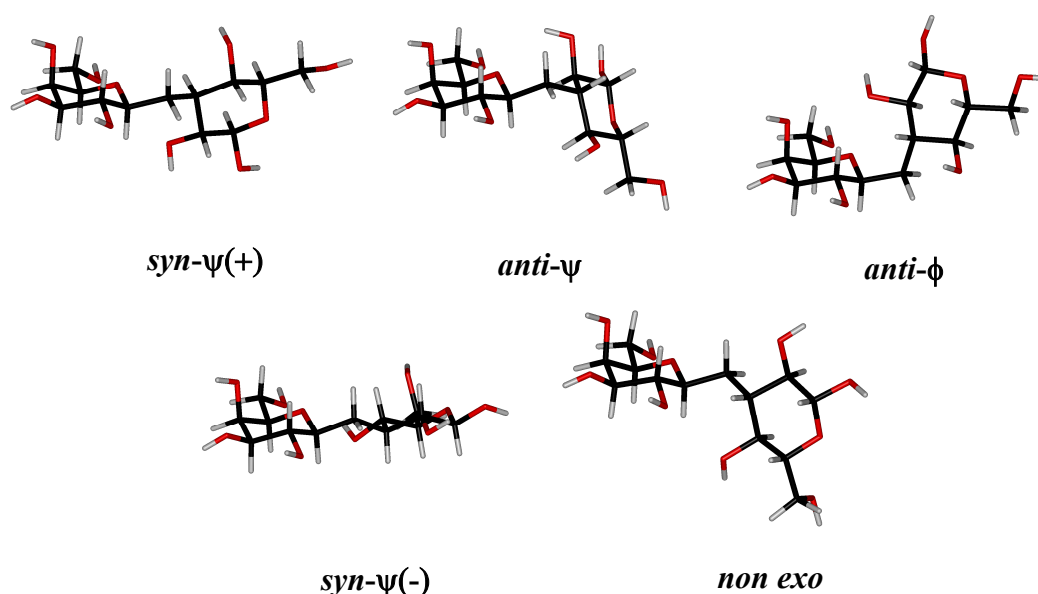


Figura 5.10: Conórmeros del C-disacárido β -C-Gal-(1 \rightarrow 3)- β -Glc-OMe.

	Estudio de esta tesis
Gal1-H7pS	1.3
Gal1-H7pR	9.8
Glc3-H7pS	5.5
Glc3-H7pR	4.1

Tabla 5.5: Constantes de acoplamiento inter-glicosídicas del β -C-Gal-(1 \rightarrow 3)- β -Glc-OMe.

Siguiendo con la misma metodología que en los dos casos anteriores se obtuvieron 189350 soluciones compatibles con las constantes de acoplamiento, siendo la mejor solución la que corresponde a una distribución *syn-ψ(+):anti-ψ:anti-φ:syn-ψ(-):non exo* de 15:66:13:6:0, con un valor de M de 0.991. Si comparamos la distribución de JAMFIS con la obtenida mediante mecánica molecular (**Tabla 5.6**) se

observa que en este caso no existe un buen acuerdo entre las poblaciones teóricas y las que se deducen del análisis de constantes de acoplamiento.

	MM3*	JAMFIS
<i>syn</i> - ψ (+)	60	15
<i>anti</i> - ψ	24	66
<i>anti</i> - ϕ	1	13
<i>syn</i> - ψ (-)	12	6
<i>non exo</i>	3	0

Tabla 5.6: Resultados de la distribución de población calculada con MM3* y la obtenida empleando el método de JAMFIS para el β -C-Gal-(1 \rightarrow 3)- β -Glc-OMe.

Los valores de las constantes de acoplamiento calculados para esta distribución se recogen en la **Figura 5.11** junto a los valores experimentales y los calculados con el campo de fuerzas MM3*.

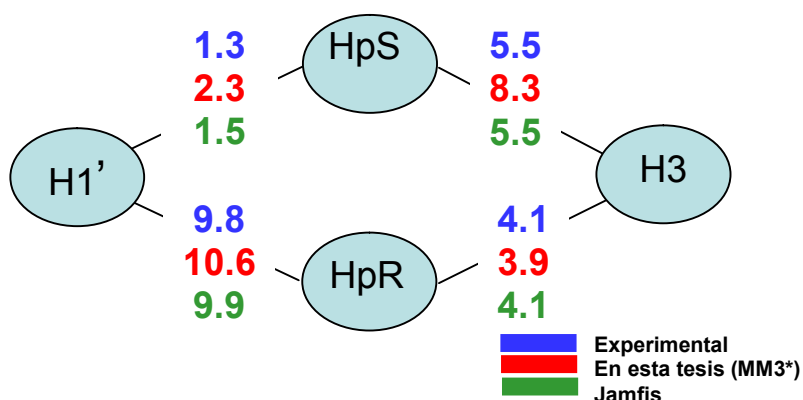


Figura 5.11: Diagrama representativo de las constantes de acoplamiento obtenidas experimentalmente, en el estudio realizado en esta tesis y aplicando NAMFIS para el β -C-Gal-(1 \rightarrow 3)- β -Glc-OMe.

El análisis de todas las soluciones posibles (**Figura 5.12**), indica que la familia *anti*- ψ está sensiblemente poblada en disolución (curva de tipo 3) aunque las constantes de acoplamiento no se pueden justificar con una única familia conformacional. Las gráficas indican que las familias conformacionales *syn*- ψ (+) y *anti*- ϕ (curvas de tipo 3 cercanas a tipo 1) estarían también presentes en disolución, aunque con una población mucho menor que la del conformero *anti*- ψ . Las curvas de los otros dos conformeros son de tipo 1 por lo que su existencia no se puede confirmar mediante constantes de acoplamiento.

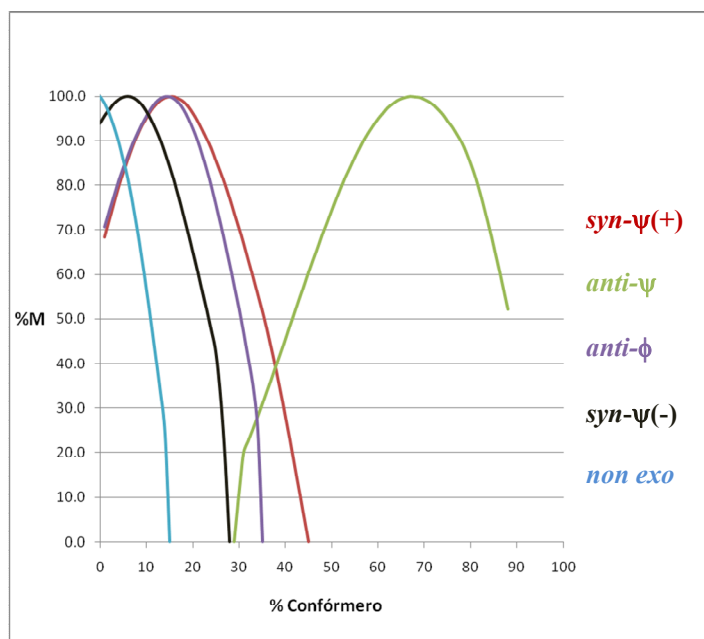


Figura 5.12: Representación del valor de M normalizado frente a la población de cada conformero del C-disacárido β -C-Gal-(1 \rightarrow 3)- β -Glc-OMe.

Si se analiza el valor de M en función del número de conformeros se obtiene la gráfica representada en la **Figura 5.13**. Con un único conformero las constantes de acoplamiento no son compatibles con las medidas deducidas experimentalmente y el valor de M es muy bajo, por lo se necesitan al menos dos conformeros para que el valor de M sea aceptable. Con tres conformeros (*anti-ψ*, *syn-ψ(+)* y *anti-φ*) se consigue un buen ajuste con un valor de M de 0.94, lo que sugeriría la presencia de tres familias conformacionales en disolución.

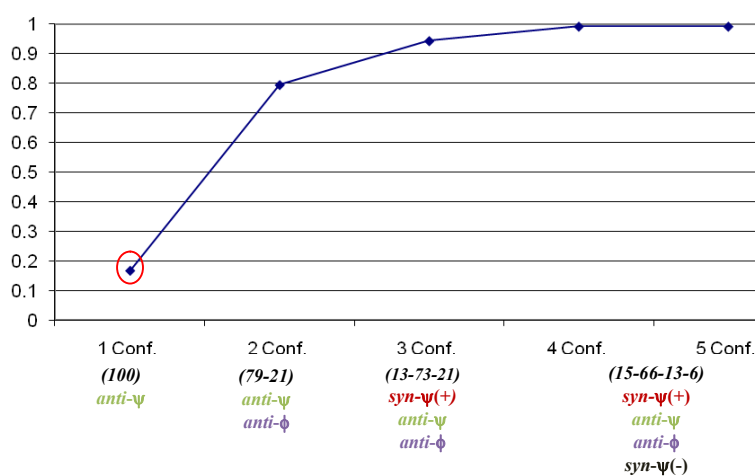


Figura 5.13: Representación del valor de M frente al número de conformeros para el β -C-Gal-(1 \rightarrow 3)- β -Glc-OMe.

Por tanto, el análisis de JAMFIS indica que la población del mínimo global *syn*- ψ (+) es menor que la calculada mediante MM3*, mientras que la del mínimo *anti*- ψ es mucho mayor que la estimada, siendo el conformero mayoritario en disolución. En el análisis previo de este C-disacárido a partir de datos de NOE y constantes de acoplamiento (capítulo 2) se llegó a una conclusión similar, aunque de forma cualitativa. Sin embargo, no fue posible confirmar la presencia del conformero *anti*- ϕ ya que el NOE exclusivo de esta conformación (Gal2'-Glc3) no pudo detectarse debido al solapamiento de señales.

4. Aplicación a C-disacáridos analizados por Kishi:

4.1 Enlace (1→1): α,α -, β,β - y α,β -C-trehalosa.

Las trehalosas son disacáridos no reductores formados por dos unidades de glucosa con uniones 1→1. Existen tres isómeros: α,α -, α,β - y β,β -trehalosa. Mientras que la α,α -trehalosa está ampliamente distribuida en la naturaleza, tanto la α,β - y β,β - son compuestos sintéticos. En el caso de derivados 1→1 ambos enlaces están afectados por el efecto *exo*-anomérico, por lo que debemos de hablar de ángulos glicosídicos ϕ y ϕ' .

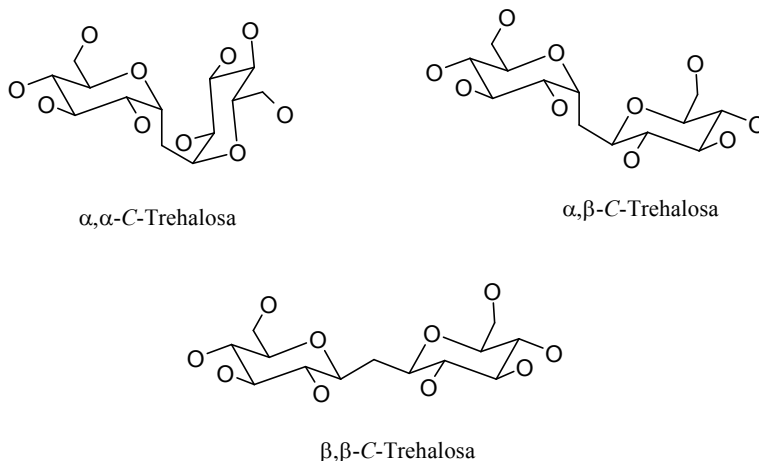


Figura 5.14: Posibles isómeros de la C-trehalosa.

Los correspondientes C-disacáridos fueron sintetizados y analizados por Kishi¹³. Las constantes de acoplamiento inter-glicosídicas están recogidas en la **Tabla 5.7**. En el caso de los C-disacáridos simétricos α,α - y β,β - las constantes de acoplamiento interglicosídicas se midieron en derivados no simétricos acetilados en O2', y se asumió que la acetilación en posición 2' no alteraba las propiedades conformacionales del C-disacárido, lo que había sido confirmado experimentalmente por Kishi en compuestos relacionados. Tras el análisis cualitativo de la constantes de acoplamiento mediante la aproximación de la matriz de diamante, Kishi concluyó que los derivados simétricos α,α - y β,β - son mono-conformacionales en disolución adoptando una disposición *exo*-anomérica en torno a ambos enlaces, mientras que en la α,β -C-trehalosa existe una preferencia por una conformación *exo*-anomérica en torno al enlace axial, mientras que en torno al ecuatorial no se pudo determinar si se

adopta una conformación distorsionada con respecto a la *exo*-anomérica o si existe un equilibrio de conformeros alternados.

	α,α -trehalosa*	α,β -trehalosa	β,β -trehalosa *
Glc1'-Hα	11.5	9.7	2.0
Glc1'-Hα'	3.5	4.0	11.1
Glc1- Hα	3.5	7.1	10.9
Glc1- Hα'	11.4	3.7	2

Tabla 5.7: Constantes de acoplamiento inter-glicosídicas reportados para los tres isómeros de la trehalosa. Los protones del puente metilénico se nombran como α y α' .

Posteriormente el grupo de S. Pérez²⁵ llevó a cabo estudios de mecánica molecular con el campo de fuerzas MM3 para estos C-disacáridos. Estos estudios predijeron que la α,α -trehalosa es rígida, con ambos enlaces adoptando una conformación *exo*-anomérica, mientras que las α,β - y β,β -trehalosas eran más flexibles y que, para ellas, además de la conformación doblemente *exo*-anomérica, que era mayoritaria, existían pequeñas poblaciones de otras familias conformacionales.

Para el estudio de estos compuesto por JAMFIS tomaremos como punto de partida los mínimos determinados por Pérez (**Figura 5.15**) junto con las constantes de acoplamiento medidas por Kishi. Por consistencia con los estudios realizados para otros C-disacáridos, las geometrías se minimizaron con el campo de fuerzas MM3* y dieléctrica 80. En la **Tabla 5.8** se incluyen las poblaciones de cada conformero con las que se obtiene el mejor ajuste con JAMFIS, junto con las calculadas a partir de la energía relativa de los conformeros minimizados con MM3*. En la **Figura 5.16** se representan las contantes promedio obtenidas con JAMFIS junto a las calculadas a partir de la distribución de MM3* y los valores experimentales.

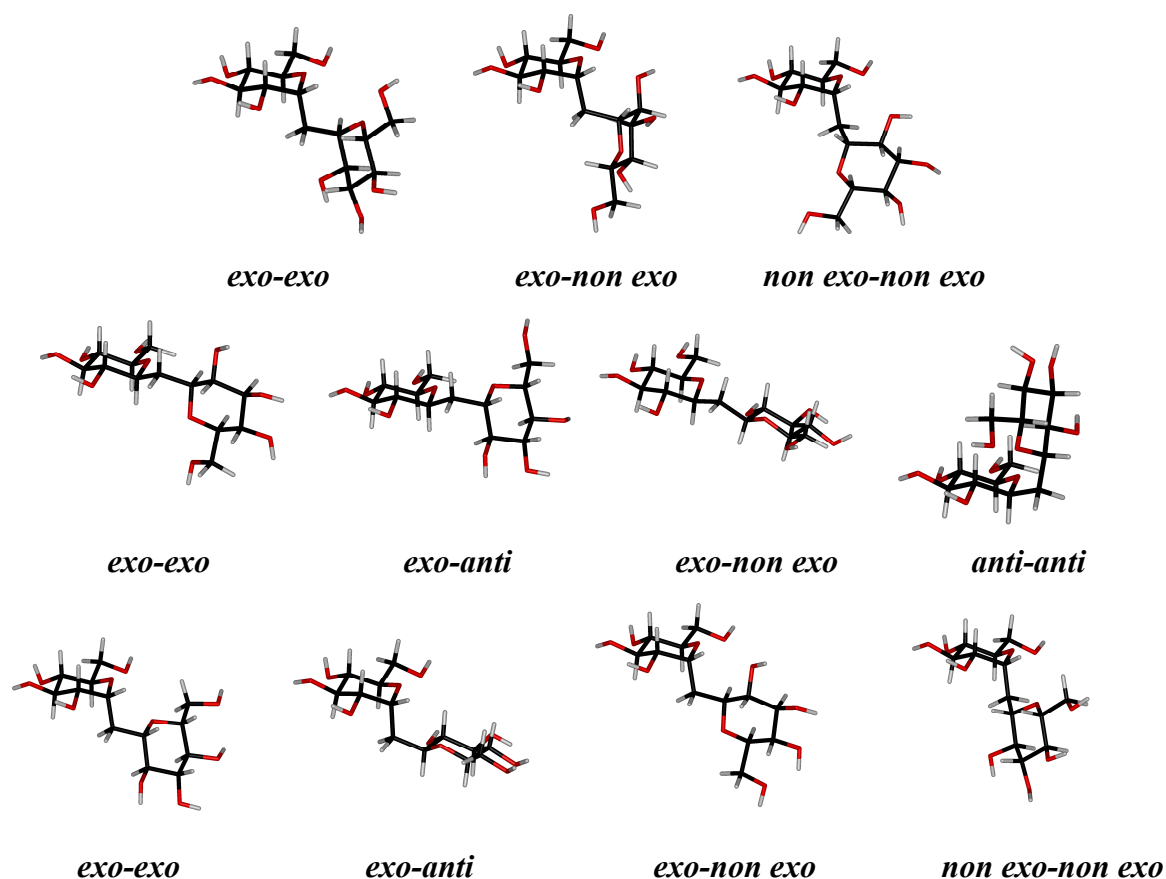


Figura 5.15: Conórmeros de las tres C-trehalosas: $\alpha\alpha$ -trehalosa(arriba), $\beta\beta$ -trehalosa (centro) y $\alpha\beta$ -trehalosa (abajo).

		MM3	MM3*	JAMFIS
$\alpha\alpha$	<i>exo-exo</i>	100	99	92
	<i>exo-non exo</i>	0	1	0
	<i>non exo-non exo</i>	0	0	8
$\alpha\beta$	<i>exo-exo</i>	82	73	32
	<i>exo-anti</i>	8	4	37
	<i>exo-non exo</i>	6	15	10
	<i>non exo-exo</i>	4	8	21
$\beta\beta$	<i>exo-exo</i>	75	83	92
	<i>exo-anti</i>	13	15	2
	<i>exo-non exo</i>	8	2	0
	<i>anti-anti</i>	4	0	6

Tabla 5.8: Distribución de población de las C-trehalosas calculadas mediante mecánica molecular (MM3* y $\epsilon = 80$ y MM3 $\epsilon = 4$) y mediante el método JAMFIS.

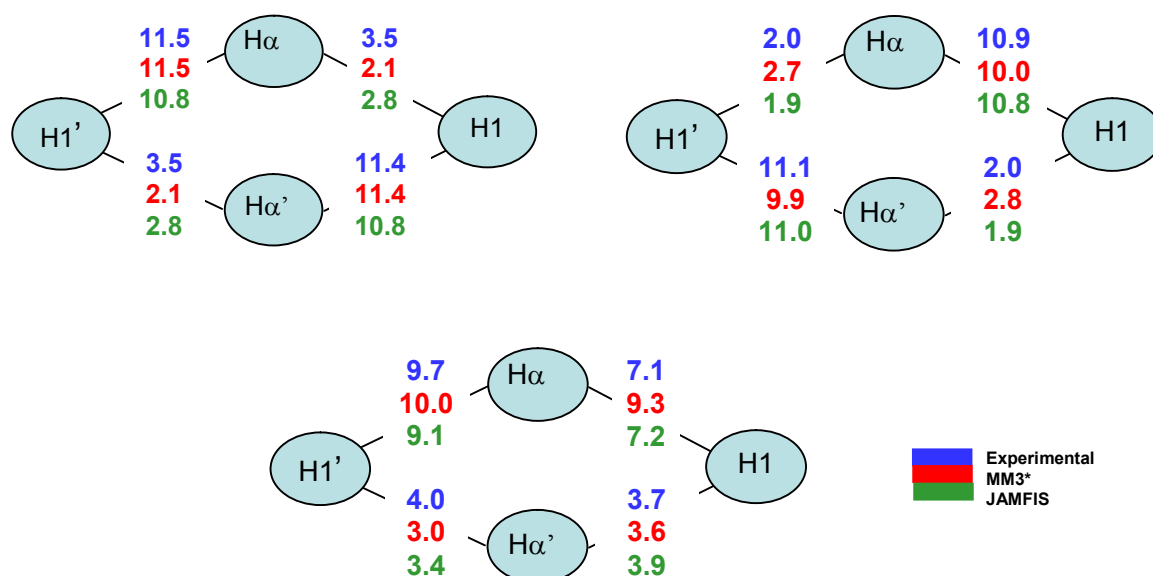


Figura 5.16: Diagrama representativo de las constantes de acoplamiento de las tres C-trehalosas: α,α - (arriba izquierda), β,β - (arriba derecha) y $\alpha\beta$ - (abajo).

Tanto en la α,α -C-trehalosa como en la β,β -C-trehalosa el mejor ajuste se obtiene cuando el conformero *exo/exo* está muy poblado (92% en ambos casos), lo que es consistente con los resultados de mecánica molecular, sobre todo en el caso de la α,α -C-trehalosa, para la que se predice un comportamiento casi monoconformacional. Sin embargo, en la α,β -C-trehalosa la distribución de población que proporciona el mejor ajuste es 32:37:10:21, correspondiente a los mínimos *exo/exo*: *exo/anti*: *exo/non exo*: *non exo/exo* respectivamente, lo que no estaría de acuerdo con los cálculos teóricos que predicen que el conformero *exo/exo* es muy mayoritario.

El análisis general de todas las soluciones (**Figura 5.17**) indica que en las C-trehalosas simétricas la curva del conformero *exo/exo* es de tipo 2, mientras que las de los otros conformeros posibles son de tipo 1, lo que confirma que el conformero *exo/exo* es muy predominante en disolución. Por el contrario, en la α,β -C-trehalosa, tres de las curvas son de tipo 3, por lo que las constantes de acoplamiento experimentales sólo se podrían explicar con la existencia de tres familias conformacionales en disolución (*exo-exo*, *exo-anti* y *non exo/exo*). En cambio, la presencia del conformero *exo-non exo*, con una curva de tipo 1, no puede confirmarse.

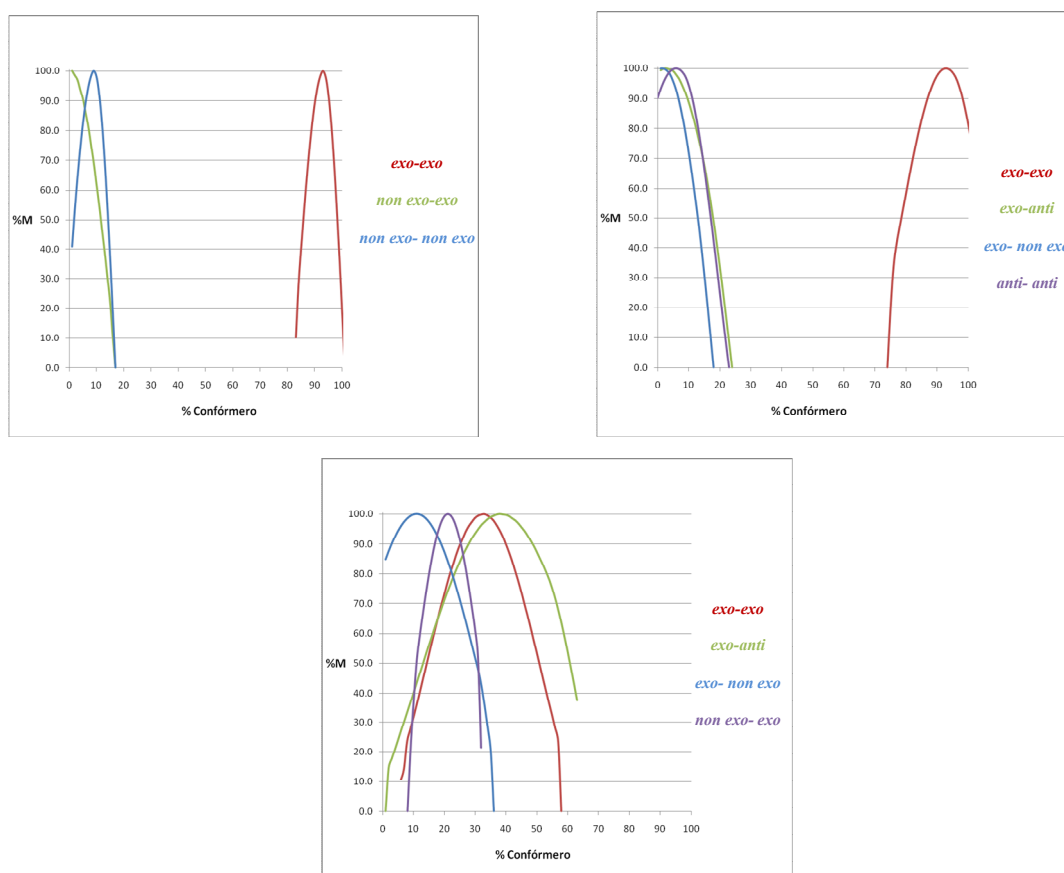


Figura 5.17: Representación del valor de M en función de la población de cada conformero para los tres isómeros de la C-trehalosa: α,α - (arriba izquierda), β,β - (arriba derecha) y α,β - (abajo).

Por último, si se analizan las gráficas obtenidas mediante el método 3 (**Figura 5.18**), en las dos C-trehalosas simétricas la diferencia entre las constantes de acoplamiento del conformero *exo/exo* y las experimentales es menor que 1.5 Hz y por tanto, la solución mono-conformacional se encuentra dentro del error establecido. No obstante, el ajuste entre los valores experimentales y calculados mejora si se considera una pequeña población de un segundo conformero (8% del conformero *non exo/non exo* en la α,α -C-trehalosa y 7% del conformero *anti/anti* en la β,β -C-trehalosa). Por otra parte, la introducción de conformeros adicionales no supone ninguna mejora en el valor de M. Este resultado contrasta con el de la α,β -C-trehalosa, para la que se necesitarían al menos tres familias conformacionales (*exo/exo*, *exo/anti*, *non exo/exo*) para obtener una solución dentro del error y con un valor de M aceptable. Si en este último caso introducimos una pequeña población de

una cuarta familia conformacional (*exo/non exo*) es posible conseguir un valor de M aún mayor.

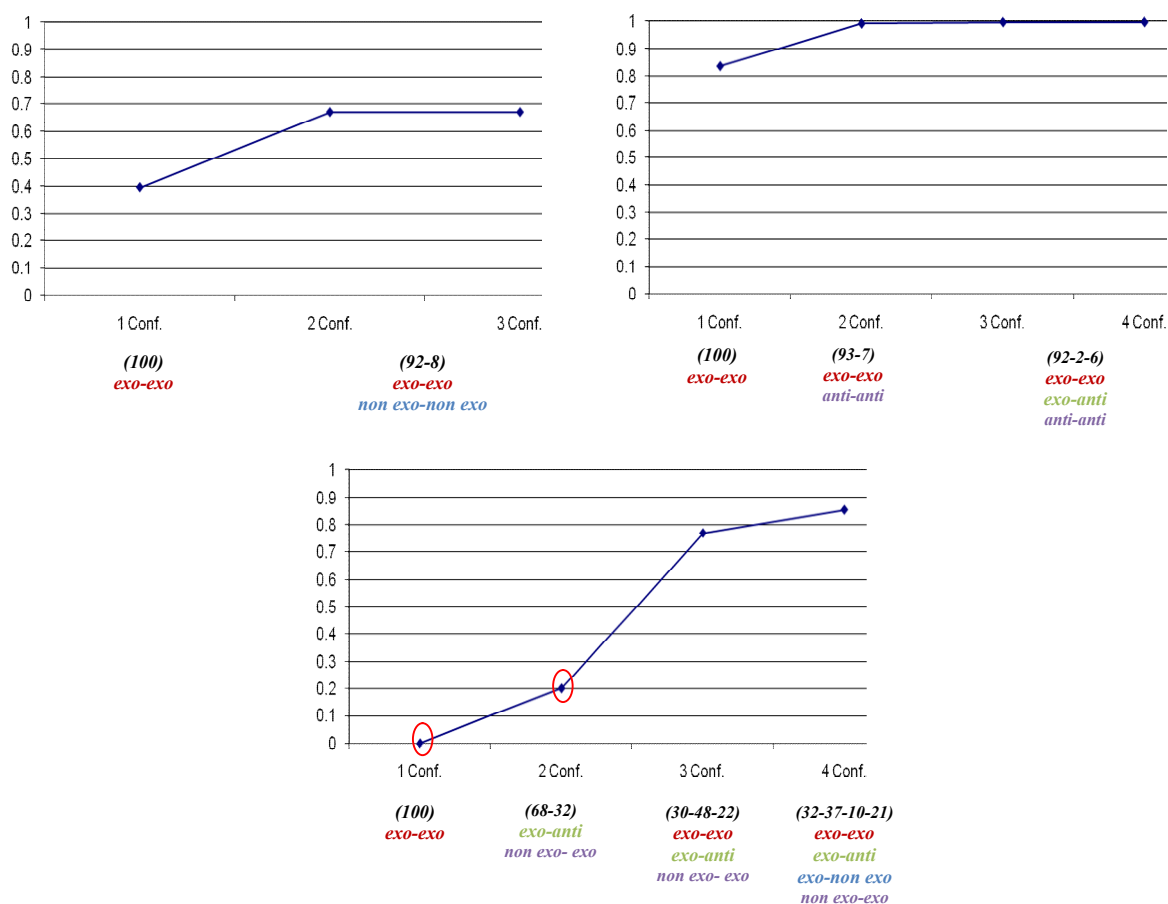


Figura 5.18: Representación del valor de M frente al número de cónformeros para los isómeros de la C-trehalosa. : α,α- (arriba izquierda), β,β- (arriba derecha) y α,β- (abajo).

Por tanto, los resultados de JAMFIS indican que la α,α- y la β,β-C-trehalosa adoptan predominantemente una conformación *exo*-anomérica en torno a ambos enlaces, lo que está de acuerdo con las conclusiones de Kishi. En el caso de la α,β-C-trehalosa, con la misma información experimental con la que contaba Kishi, el método JAMFIS ha permitido caracterizar el comportamiento conformacional con mucho más detalle y se ha demostrado la flexibilidad de este isómero en torno a ambos enlaces. El análisis de Kishi indicó la preferencia por una disposición *exo*-anomérica en torno al enlace axial, mientras que en torno al enlace ecuatorial no se pudo distinguir entre la existencia de una conformación distorsionada respecto a la *exo*-anomérica o de un equilibrio de conformaciones alternadas. Sin embargo, el análisis JAMFIS indica, que aunque la conformación *exo*-anomérica en torno al

enlace axial es mayoritaria, el conformero *non exo* también está poblado de manera significativa y que, en torno al enlace ecuatorial, existe un equilibrio conformacional entre los conformeros *exo* y *anti*.

4.2 C-Disacáridos con uniones $\alpha(1\rightarrow3)$:

Existen tres C-disacáridos con unión $\alpha(1\rightarrow3)$, que son epímeros en las posiciones 2 ó 4 (**Figura 5.19**), que fueron también analizados por Kishi. Dos de ellos, α -C-Gal-(1 \rightarrow 3)- α -Gal-OMe⁴ y α -C-Gal- $\alpha(1\rightarrow3)$ - α -Glc-OMe⁵, son los C-análogos de los disacáridos constituyentes de los determinantes sanguíneos trisacárido A y trisacárido B, mientras que el α -C-Gal(1 \rightarrow 3)- α -Man-OMe se sintetizó en el grupo de Vogel²⁶. Las constantes de acoplamiento medidas para estos disacáridos se recogen en la **Tabla 5.9**. A partir del análisis basado en la matriz de diamante, Kishi concluye que los tres C-disacáridos adoptan la conformación *exo*-anomérica en torno al enlace C-glicosídico, mientras que en torno al enlace C-aglicónico presentan diferencias significativas, lo que pone de manifiesto la importancia de la orientación axial o ecuatorial del hidroxilo contiguo al enlace aglicónico en el comportamiento conformacional. Según Kishi la conformación predominante en torno a este enlace sería la *syn*(-) en el C-Gal-Gal, la *syn*(+) en el C-Gal-Man, mientras que no existiría una conformación definida en el C-Gal-Glc.

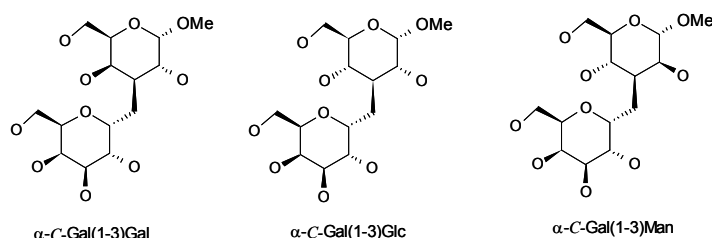


Figura 5.19: Estructuras α -C-Gal(1 \rightarrow 3)- α -Gal-OMe, α -C-Gal(1 \rightarrow 3)- α -Glc-OMe y α -C-Gal(1 \rightarrow 3)- α -Man-OMe.

	α -C-Gal(1 \rightarrow 3)- α -Gal-OMe	α -C-Gal(1 \rightarrow 3)- α -Glc-OMe	α -C-Gal(1 \rightarrow 3)- α -Man-OMe
Gal1-H7pS	12.1	11.0	10.2
Gal1-H7pR	3.1	3.3	2.4
Gal3/Glc3-H7pS	3.3	3.6	9.2
Gal3/Glc3-H7pR	11.1	6.1	4.2

Tabla 5.9: Constantes de acoplamiento inter-glicosídicas medidas por Kishi y Vogel para C-disacáridos con unión $\alpha(1\rightarrow3)$.

Dado que en este caso no existen estudios computacionales para estos tres C-disacáridos, el primer paso fue realizar una búsqueda conformacional sistemática mediante mecánica molecular con el campo de fuerzas MM3* para determinar los

conformeros energéticamente posibles. Se encontraron 4 posibles mínimos que se representan en la **Figura 5.20** para el α -C-Gal(1 \rightarrow 3)- α -Gal-OMe. Estos conformeros, junto con las constantes de acoplamiento interglucosídicas medidas por Kishi y Vogel (**Tabla 5.9**) fueron el punto de partida del estudio de JAMFIS.

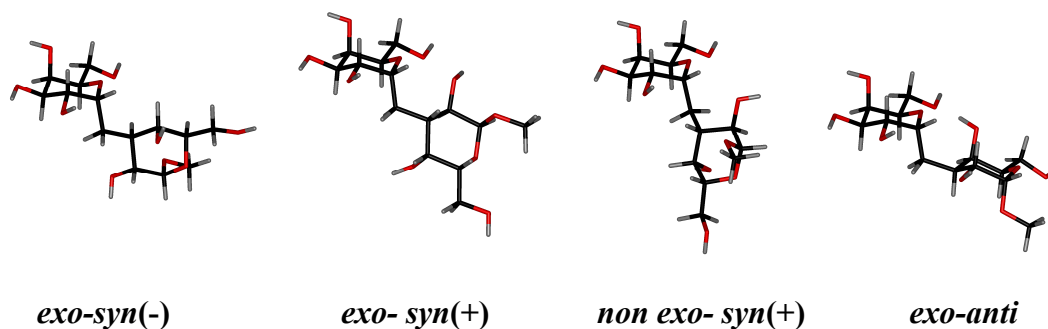


Figura 5.20: Conformeros de la α -C-Gal-(1 \rightarrow 3)- α -Gal-OMe.

Tras la aplicación de JAMFIS se obtuvieron las poblaciones que se incluyen en la **Tabla 5.10** junto a las calculadas mediante MM3* y las constantes promedio que se representan en la **Figura 5.21** junto a los valores experimentales y calculados con MM3*. Por otro lado en la **Figura 5.22** se representan las curvas completas del valor de M en función de la población de cada conformero y en **Figura 5.23** la variación de M con el número de conformeros.

		MM3*	NAMFIS
α -C-Gal-(1 \rightarrow 3)- α -Gal-OMe	<i>exo-syn(-)</i>	97	89
	<i>exo-syn(+)</i>	1	6
	<i>non exo-syn(+)</i>	1	1
	<i>exo-anti</i>	1	4
α -C-Gal-(1 \rightarrow 3)- α -Glc-OMe	<i>exo-syn(-)</i>	82	33
	<i>exo-syn(+)</i>	5	0
	<i>non exo-syn(+)</i>	4	10
	<i>exo-anti</i>	9	57
α -C-Gal-(1 \rightarrow 3)- α -Man-OMe	<i>exo-syn(-)</i>	51	18
	<i>exo-syn(+)</i>	36	58
	<i>non exo-syn(+)</i>	12	9
	<i>exo-anti</i>	1	15

Tabla 5.10: Distribución de población obtenida mediante cálculos de mecánica molecular (MM3* y $\epsilon = 80$) y mediante el método JAMFIS para los C-disacáridos con unión α -(1 \rightarrow 3)

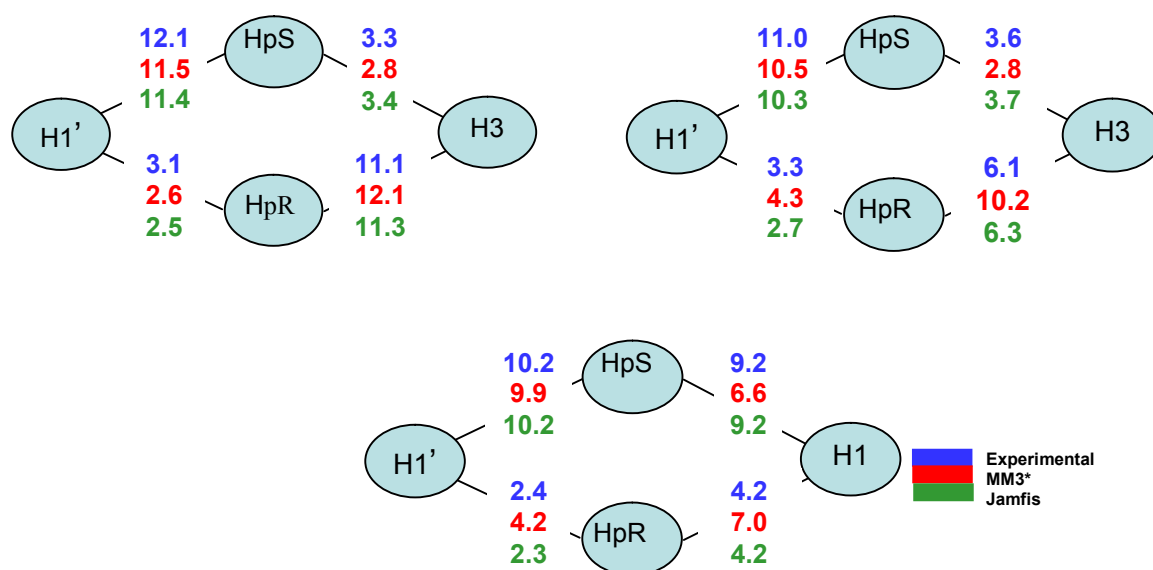


Figura 5.21: Diagrama representativo de las constantes de acoplamiento de las $\alpha(1\rightarrow3)$ derivados: α -C-Gal(1 \rightarrow 3)-Gal (arriba izquierda, α -C-Gal(1 \rightarrow 3)-Glc (arriba derecha) y α -C-Gal(1 \rightarrow 3)-Man (abajo).

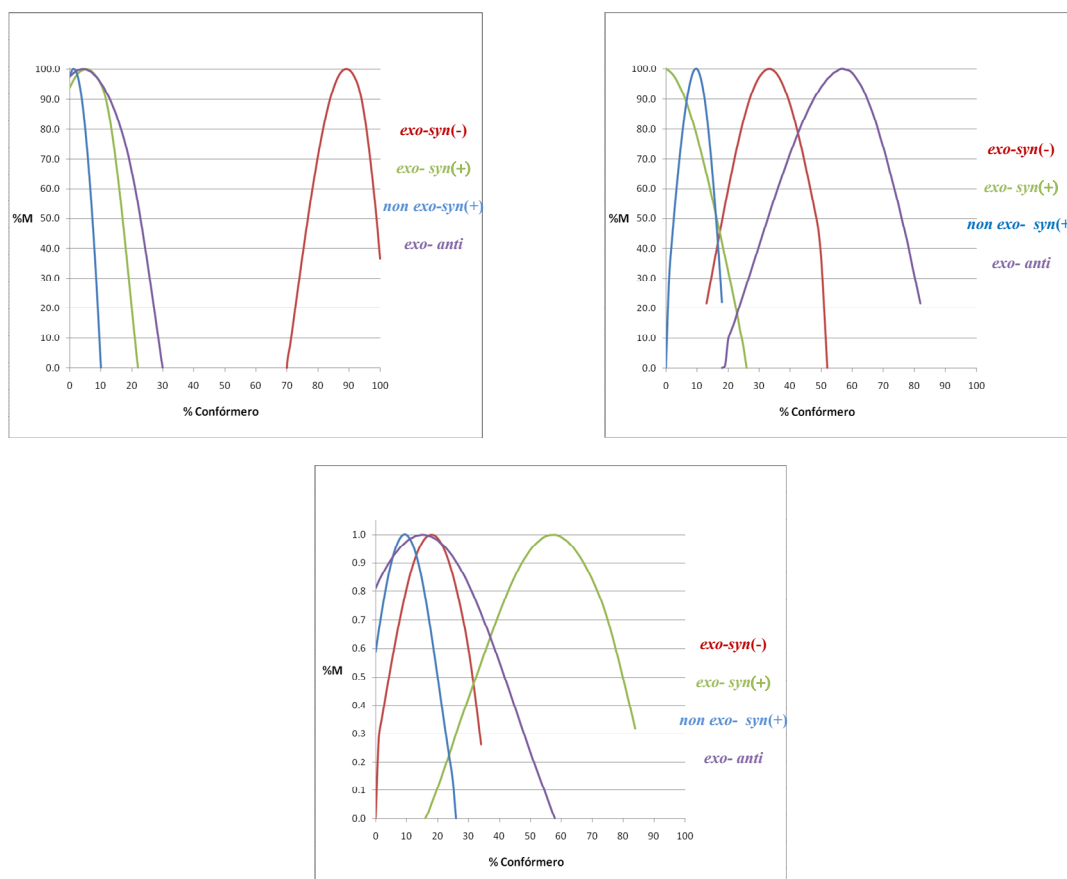


Figura 5.22: Representación del valor de M en función de la población de cada conformero para los tres C-disacáridos con enlaces $\alpha(1\rightarrow3)$: α -C-Gal-(1 \rightarrow 3)- α -Gal (arriba izquierda), α -C-Gal-(1 \rightarrow 3)- α -Glc (arriba derecha) y α -C-Gal-(1 \rightarrow 3)- α -Man (abajo).

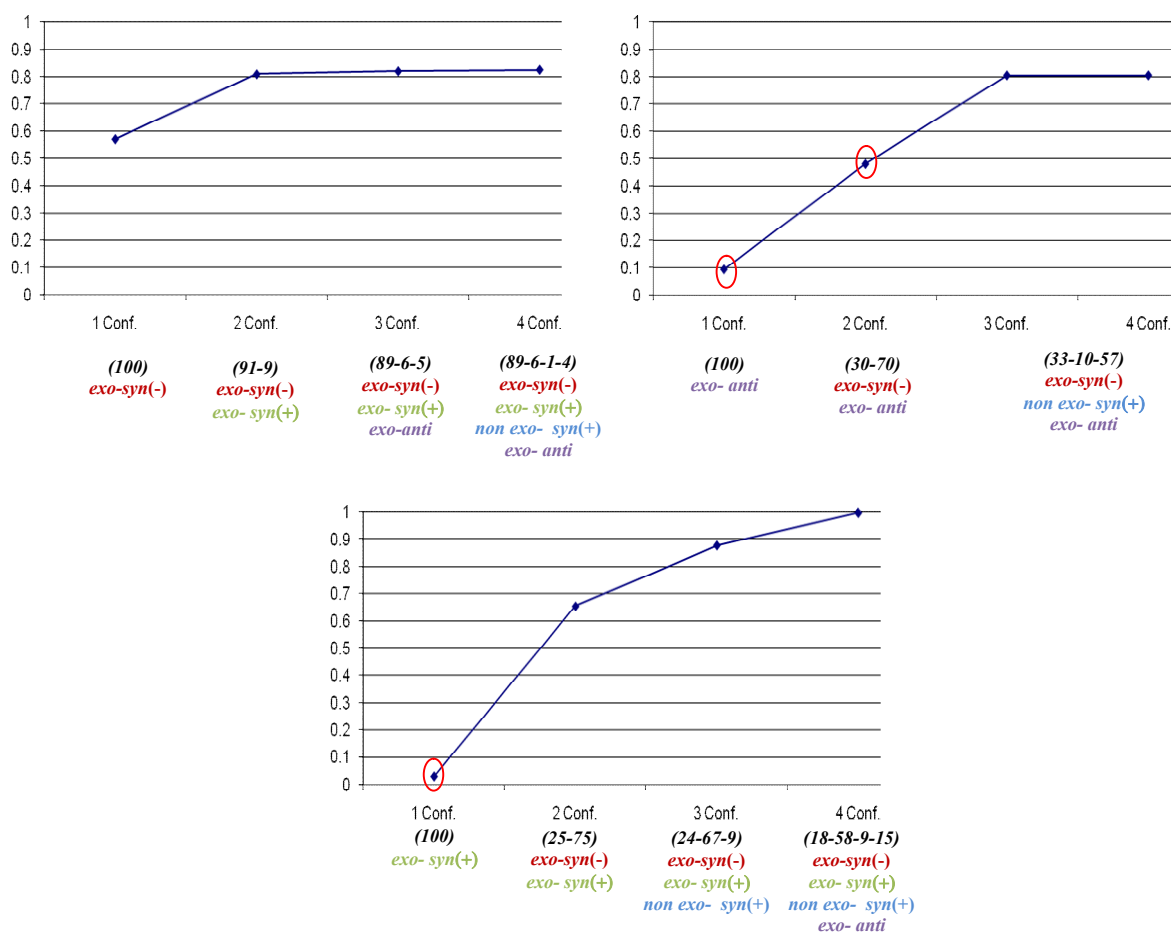


Figura 5.23: Representación del valor de M frente al número de conformeros para los tres C-disacáridos con enlaces $\alpha(1\rightarrow3)$: α -C-Gal-(1 \rightarrow 3)- α -Gal (arriba izquierda), α -C-Gal-(1 \rightarrow 3)- α -Glc (arriba derecha) y α -C-Gal-(1 \rightarrow 3)- α -Man (abajo).

α -C-Gal-(1 \rightarrow 3)- α -Gal-OMe: En la solución que proporciona el mejor ajuste el conformero *exo/syn(-)* es muy mayoritario, con una población del 89%. En la gráfica de la **Figura 5.22** puede observarse que la curva de este conformero es de tipo 2, mientras que las de los otros conformeros son de tipo 1, confirmando la predominancia de la familia conformacional *exo/syn(-)*. Este resultado está de acuerdo con los cálculos de mecánica molecular. Aunque las constantes del conformero *exo/syn(-)* son consistentes con las experimentales dentro del error de 1.5 Hz, en la **Figura 5.23** puede observarse que la introducción de una pequeña población (9%) del conformero *exo/syn(+)* permite mejorar el ajuste de forma significativa, mientras que la consideración tercer conformero ya prácticamente no aumenta el valor de M.

α -C-Gal-(1→3)- α -Glc-OMe: La mejor solución de JAMFIS es la distribución de población 33:10:57 correspondiente a las conformaciones *exo-syn*(-), *non exo-syn*(+) y *exo-anti* respectivamente, lo que no está de acuerdo con los cálculos de MM3* que predicen que la conformación *exo-syn*(-) es predominante. El análisis de la gráfica de la **Figura 5.22** confirma que los conformeros *exo/syn*(-) y *exo-anti* presentan curvas de tipo 3, por lo que ambos deben estar presentes en disolución para poder explicar las constantes de acoplamiento. La curva del conformero *non exo-syn*(+) es también de tipo 3, aunque cercana a tipo 1, indicando la existencia de una pequeña población de esta familia conformacional. Este resultado se confirma con la gráfica de la **Figura 5.23** que revela que es necesario tener en cuenta 3 conformeros para encontrar una distribución de población que sea compatible con las constantes de acoplamiento experimentales.

α -C-Gal-(1→3)- α -Man-OMe: La mejor solución de JAMFIS es 18:58:9:15 correspondiente a las conformaciones *exo-syn*(-), *exo-syn*(+) *non exo-syn*(+) y *exo-anti*, lo que sugiere que varias conformaciones podrían estar pobladas en disolución. El análisis de las gráficas confirma que hay un conformero, *exo/syn*(+), con una curva de tipo 3 bastante clara, mientras que el conformero *exo-syn*(-) también tiene una curva de tipo 3, aunque más cercana a tipo 1. Por tanto, ambos conformeros estarían presentes en disolución, aunque la población del conformero *exo/syn*(+) sería mucho mayor. Por otro lado, los conformeros *non exo-syn* y *exo-anti* presentan curvas de tipo 1, especialmente ancha para este último conformero, por lo que su presencia no puede confirmarse de forma concluyente. La gráfica de la **Figura 5.23** si parece sugerir que estos dos conformeros podrían estar presentes ya que el valor de M pasa de 0.65 a 0.89 cuando se tiene en cuenta un tercer conformero y a un valor de 0.99 si se incluye un cuarto, por lo que el ajuste mejora de forma considerable con respecto a la solución con sólo dos conformeros.

Por tanto, el análisis de JAMFIS indica que el isómero con las dos unidades de galactosa adopta predominantemente una conformación *exo-syn*(-), lo que está de acuerdo con las conclusiones de Kishi. Por el contrario, cuando la unidad de galactosa es reemplazada por glucosa o manosa, los C-disacáridos resultantes son más flexibles y presentan más de una familia conformacional en disolución. Así, en el isómero α -C-Gal-(1→3)- α -Glc-OMe, estarían poblados tres conformeros, *exo*-

syn(-), *non exo-syn(+)* y *exo-anti*, siendo éste último el mayoritario. En este C-disacárido, Kishi concluyó que no existe una conformación preferida en torno al enlace aglicónico lo que está de acuerdo con este resultado, pero no le fue posible definir los conformeros poblados, ni detectar la presencia de la conformación minoritaria "no natural" *non exo* en torno al enlace glicosídico. Finalmente, en el isómero α -C-Gal-(1 \rightarrow 3)- α -Man-OMe, se necesita la presencia de los conformeros *exo-syn(+)* y *exo-syn(-)* para justificar los valores de las constantes de acoplamiento, lo que pone de manifiesto la movilidad del C-disacárido en torno al enlace aglicónico. En este caso Kishi concluyó que el conformero *exo-syn(+)* es predominante, pero no fue capaz de determinar la presencia de ninguna familia conformacional adicional.

4.3 C-Disacáridos con uniones $\alpha(1\rightarrow4)$

Varios C-disacáridos con uniones $\alpha(1\rightarrow4)$ fueron sintetizados y estudiados por el grupo de Kishi¹⁰ (**Figura 5.24**). El análisis del C-maltósido (α -C-Glc-(1 \rightarrow 4)- α -Glc-OMe), el C2'-epi-C-maltósido (α -C-Man-(1 \rightarrow 4)- α -Glc-OMe), y el C3-deoxi-C-maltósido permitió investigar el comportamiento conformacional de la unión α -C-(1 \rightarrow 4), así como determinar la influencia de la disposición axial o ecuatorial del hidroxilo en C2' y de la presencia o ausencia del hidroxilo en C3.

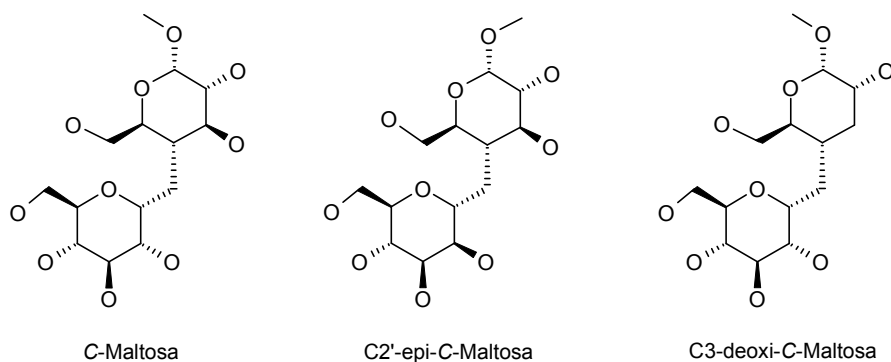


Figura 5.24: Estructuras de C-maltosa, C2'-epi-C-maltosa y C3-deoxi-C-maltosa.

Las conclusiones del análisis de Kishi fueron que la conformación *exo*-anomérica en torno al enlace glicosídico es predominante, independientemente de la configuración de C2' y que la existencia de una interacción estérica desestabilizante tipo 1,3-diaxial en cada una de las conformaciones alternadas en torno al enlace aglicónico evita que el C-disacárido adopte una conformación mayoritaria en torno a este enlace, salvo en el caso del derivado 3-desoxi, en el que el conformero *syn*(-), que no tiene ninguna interacción de este tipo, es predominante.

En el caso de la C-maltosa existen estudios previos de mecánica molecular con el campo de fuerzas MM3 del grupo de Pérez²⁷ y los tres mínimos encontrados se utilizaron como geometrías de partida para una minimización posterior con MM3* y una constante dieléctrica de 80 debyes. En el caso de los otros dos C-disacáridos α -(1 \rightarrow 4) se realizó una búsqueda sistemática con este campo de fuerzas con el objeto de encontrar los mínimos posibles. Por otro lado, como datos experimentales, se usaron las constantes de acoplamiento escalares vecinales medidas por el grupo de Kishi para estos tres compuestos (**Tabla 5.11**).

	C-maltosa	C2'-epi-C-maltosa	C3-deoxi-C-maltosa
H1'-H7pS	10.3	9.0	12.1
H1'-H7pR	3.1	4.6	3.1
H4-H7pS	2.9	3.3	2.9
H4-H7pR	5.5	5.3	11.4

Tabla 5.11: Tabla con los valores de las constantes de acoplamiento experimentales medidas por Kishi para los tres C-disacáridos con unión $\alpha(1\rightarrow4)$.

Tras la aplicación de JAMFIS se obtuvieron las poblaciones que se recogen en la **Tabla 5.12** junto a las calculadas mediante el campo de fuerzas MM3*. En la **Figura 5.25** se representan las constantes de acoplamiento de la distribución que proporciona el mejor ajuste, junto a los valores experimentales y a los que se predicen mediante los cálculos de mecánica molecular. Por otro lado, en la **Figura 5.26** se representan las curvas del valor de M en función de la población de cada conformero y en la **Figura 5.27** el valor de M en función del número de conformeros.

		MM3*	NAMFIS
C-Maltosa	<i>exo-syn(-)</i>	74	20
	<i>exo-anti</i>	24	68
	<i>non exo-syn(+)</i>	2	12
C2'-epi-C-maltosa	<i>exo-syn(-)</i>	65	19
	<i>exo-anti</i>	27	53
	<i>non exo-syn(+)</i>	6	17
	<i>non exo-anti</i>	2	11
C3-deoxi-C-maltosa	<i>exo-syn(-)</i>	95	90
	<i>exo-anti</i>	4	9
	<i>non exo-syn(+)</i>	1	1

Tabla 5.12: Distribución de población calculada (MM3* y $\epsilon=80$) y obtenida mediante el método JAMFIS para los C-disacáridos con unión $\alpha(1\rightarrow4)$.

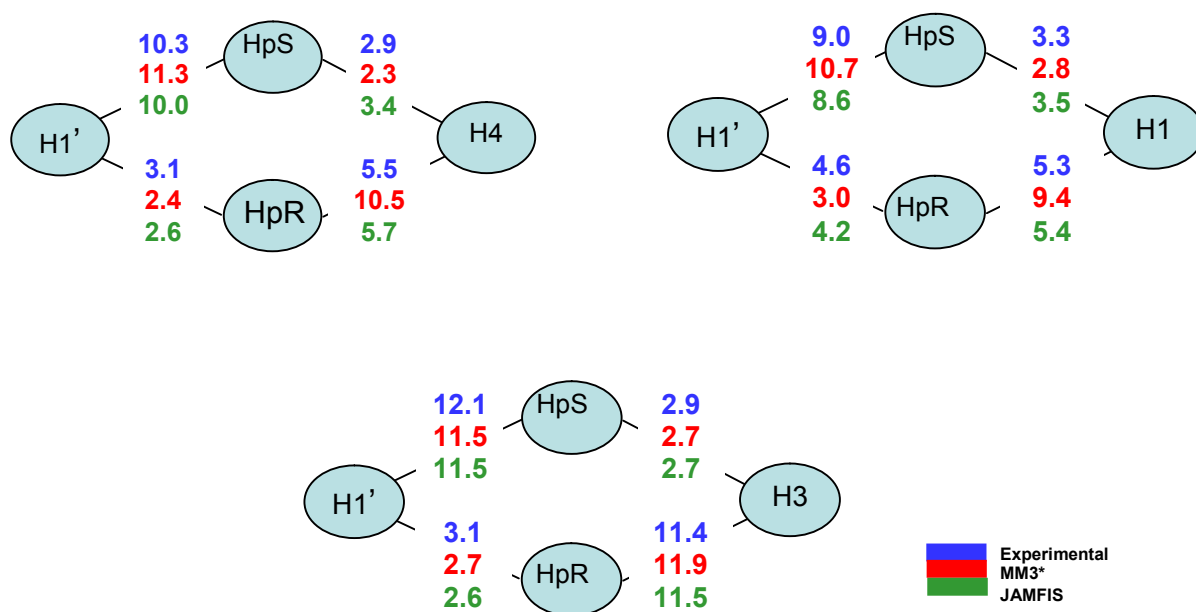


Figura 5.25 Valores de las constantes de acoplamiento de los C-disacáridos con unión $\alpha(1\rightarrow4)$: C-maltosa (arriba izquierda), C2'-epi-C-maltosa (arriba derecha) y C3-deoxi-C-maltosa (abajo).

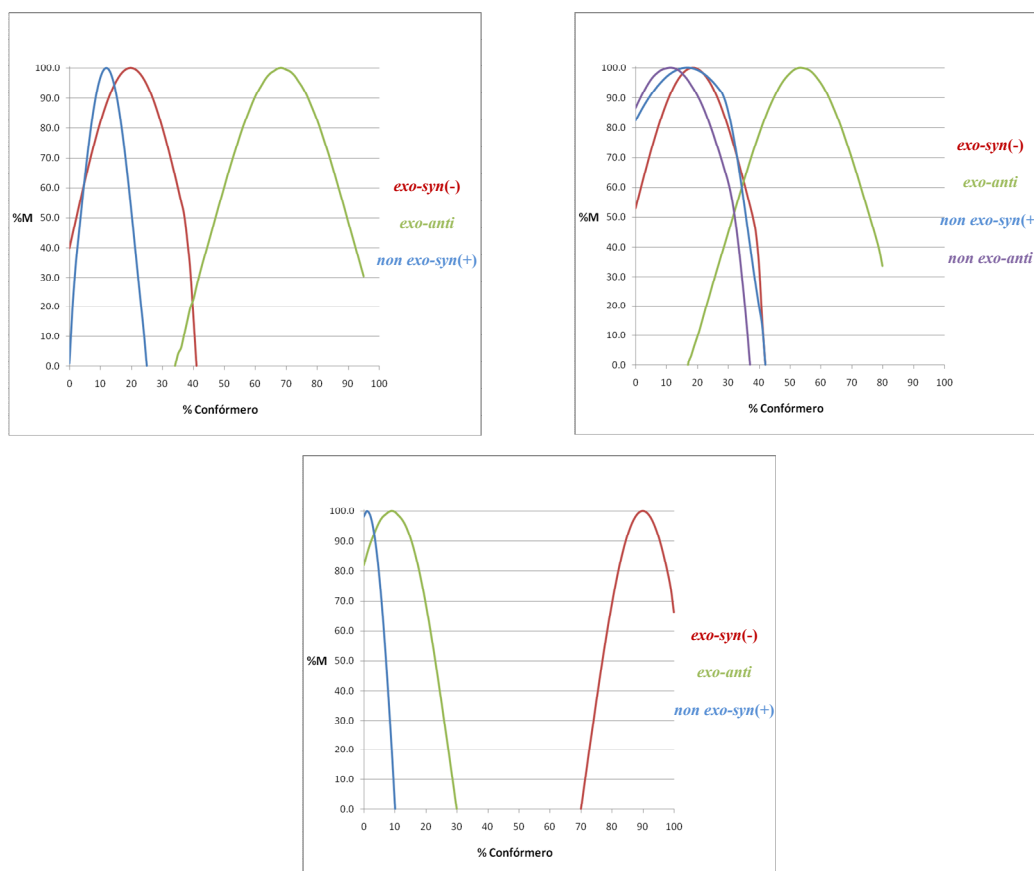


Figura 5.26: Representación del valor de M en función de la población de cada conformero para los C-disacáridos con unión $\alpha(1\rightarrow4)$: C-maltosa (arriba izquierda), C2'-epi-C-maltosa (arriba derecha) y C3-deoxi-C-maltosa (abajo).

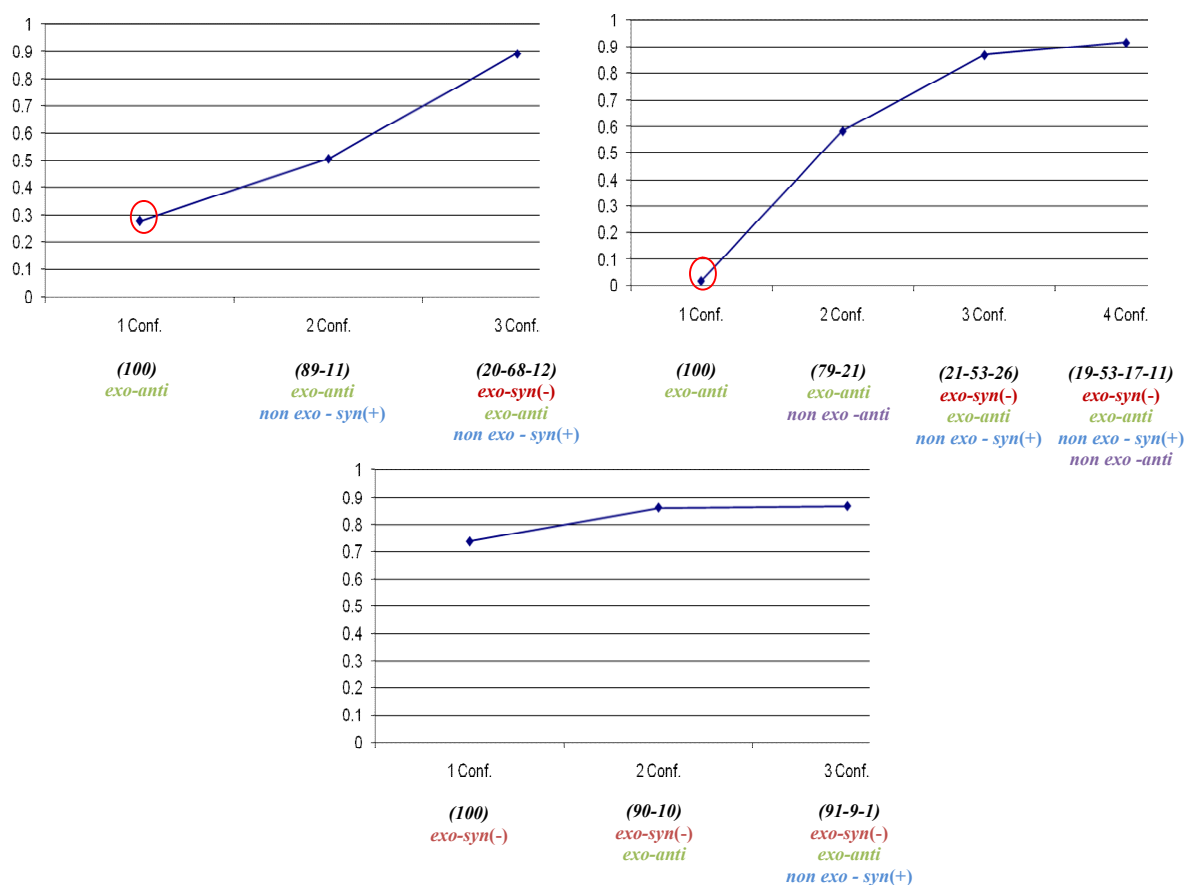


Figura 5.27: Representación del valor de M frente al número de conformeros para los C-disacáridos con unión $\alpha(1\rightarrow4)$: C-maltosa(arriba izquierda), C2'-epi-C-maltosa (arriba derecha) y C3-deoxy-C-maltosa (abajo).

C-Maltosa: La solución que mejor se ajusta a los datos experimentales es una distribución 20:68:12 entre las conformaciones *exo/syn(-)*, *exo/anti*, *non exo/syn(+)*, lo que contrasta con los cálculos teóricos, que predicen que el conformero *exo/syn(-)* es el mayoritario. Las curvas obtenidas para los tres conformeros son de tipo 3, lo que significa que las tres familias conformacionales tienen que estar presentes en disolución para poder explicar los valores de las constantes de acoplamiento, siendo el conformero *exo/anti* el más poblado (**Figura 5.26**). De igual manera, si se representa el valor de M en función del número de conformeros, se observa que el ajuste es mucho mejor cuando se tienen en cuenta los tres conformeros con respecto al obtenido con uno o dos conformeros. (**Figura 5.27**).

C2'-epi-C-maltosa: En el caso del C2'-epi-C-maltosa la mejor solución de JAMFIS es 19:53:17:11 entre las geometrías *exo-syn(-)*:*exo-anti*:*non exo-syn(+)*:*non exo-anti*. Al igual que en el caso anterior, la conformación *exo-anti* es la mayoritaria en disolución, en contra con las predicciones teóricas. Las curvas de los conformeros

exo-syn (-) y *exo-anti* son de tipo 3 y similares a las obtenidas para la C-maltosa, mientras que las curvas de los confórmeros *non exo-syn*(+) y *non exo-anti* son de tipo 1, por lo que sólo puede confirmarse la presencia de las dos primeras familias conformacionales (**Figura 5.26**). Por otro lado, si analizamos la gráfica en la que se representa el valor de M frente al número de confórmeros utilizados en el cálculo, vemos que cuando se tienen en cuenta 3 confórmeros se obtiene un valor de M muy superior al obtenido con 2 confórmeros, pero no muy diferente al obtenido con cuatro (**Figura 5.27**). Esta gráfica sugiere que es necesaria la participación de al menos tres confórmeros para obtener un buen ajuste entre las constantes de acoplamiento experimentales y calculadas.

C3-deoxi-C-maltosa: En el caso de este compuesto la mejor solución de JAMFIS corresponde a una distribución 90:9:1 para los confórmeros *exo-syn*(-):*exo-anti*:*non exo-syn*(+), que es consistente con la prevista por los cálculos de mecánica molecular. Una de las curvas es de tipo 2, mientras que las otras dos son de tipo 1, por lo que sólo con el confórmero *exo-syn*(-) sería posible explicar las constantes de acoplamiento de manera muy satisfactoria (**Figura 5.26**). No obstante, cuando se considera un segundo confórmero, el ajuste mejora ligeramente, como puede verse en la gráfica de la **Figura 5.27**.

Por tanto, el análisis de JAMFIS indica que la C-maltosa es flexible tanto en torno al enlace glicosídico (conformaciones *exo* y *non exo*), como en torno al aglicónico (conformaciones *syn* y *anti*). El confórmero *exo-anti* es el mayoritario en disolución, lo que implica la existencia de diferencias conformacionales importantes con respecto al disacárido natural, en el que el confórmero *exo-syn*(-) es muy predominante. Al igual que la C-maltosa, el epímero en C2' es flexible, al menos en torno al enlace aglicónico (conformaciones *syn* y *anti*). Aunque el confórmero *exo-anti* sigue siendo el mayoritario, el cambio de orientación del hidroxilo en C2' provoca la aparición de una proporción mayor de confórmeros *non exo*-anoméricos. Por otra parte, en la C3-deoximaltosa, la ausencia del hidroxilo en C3 produce un cambio conformacional y este derivado presenta una preferencia conformacional clara tanto en torno al enlace glicosídico como alrededor del aglicónico, siendo el confórmero *exo/syn*(-) muy predominante en disolución.

4.4 C-Disacáridos con uniones $\beta(1\rightarrow4)$:

De la misma forma, otros C-disacáridos con uniones $\beta(1\rightarrow4)$ también han sido sintetizados y estudiados por el grupo de Kishi¹⁰. Entre ellos, cabe destacar la C-celobiosa (β -C-Glc-(1 \rightarrow 4)- α -Glc-OMe), el epímero en C2' (β -C-Man-(1 \rightarrow 4)- α -Glc-OMe) y el análogo de xilosa (β -C-Glc-(1 \rightarrow 4)- α -Xyl-OMe) (**Figura 5.28**). Su análisis permitió investigar las preferencias conformacionales de la unión β -C-(1 \rightarrow 4), el efecto del cambio de configuración en C2' y la ausencia del hidroximetilo en C5, respectivamente.

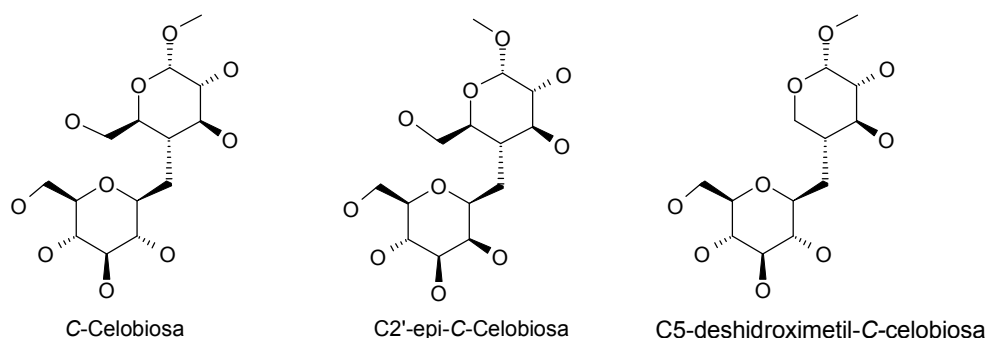


Figura 5.28: Estructuras de C-celobiosa, C2'-epi-C-celobiosa y C- β -Glc(1 \rightarrow 4)Xyl.

A partir del análisis de las constantes de acoplamiento, el grupo de Harvard concluyó que todos ellos adoptan una conformación *exo*-anomérica en torno al enlace glicosídico, como los correspondientes O-disacáridos. Por otro lado, se sugirió que el comportamiento conformacional en torno al enlace aglicónico era similar en la C-celobiosa y en su epímero en C2', pero diferente al del derivado sin el grupo hidroximetilo. Según estos autores, en los dos primeros compuestos, ninguna de las conformaciones alternadas alrededor del enlace aglicónico se encuentra libre de impedimentos estéricos, por lo que estos C-disacáridos podrían adoptar bien una conformación mayoritaria, aunque desviada de su posición alternada ideal, o bien podrían existir como una mezcla de conformeros alternados. Sin embargo, en el caso del tercer C-disacárido, postularon que existía una conformación predominante en torno a este enlace, la *exo-syn*(+), que está libre de interacciones de tipo 1,3-diaxial debido a la eliminación del grupo hidroximetilo.

La C-celobiosa ha sido estudiada previamente mediante cálculos de mecánica molecular con el campo de fuerzas MM3, en el grupo de S. Pérez²⁷, presentando los mismos mínimos que los determinados por nuestro grupo de investigación para la C-lactosa, el epímero en C4'. Para los otros dos C-disacáridos $\beta(1\rightarrow4)$ se realizó una

búsqueda sistemática y se encontraron los mismos mínimos para la C5-deshidroximetil-C-celobiosa. Sin embargo, en la C2'-epi-C-celobiosa se encontró el mínimo *non exo-anti- ψ* en lugar del *anti- ϕ* . Como en casos anteriores, estos confórmeros, junto a las constantes de acoplamiento experimentales medidas en el grupo de Y. Kishi (Tabla 5.15), se utilizaron para la aplicación de JAMFIS. La distribución de población que proporciona el mejor ajuste para cada C-disacárido se recoge en la Tabla 5.16 junto a la distribución que predice el campo de fuerzas MM3*. Los resultados del análisis se representan en la Figura 5.29 (valores de las constantes de acoplamiento), Figura 5.30 (M frente a población de cada confórmero) y Figura 5.31 (valor de M en función del número de confórmeros).

	C-celobiosa	C2'-epi-C-celobiosa	C5-deshidroximetil-C-celobiosa
H1'-H7pS	1.7	3.8	2.0
H1'-H7pR	9.2	9.1	10.0
H4-H7pS	4.7	3.8	9.5
H4-H7pR	3.5	5.2	3.1

Tabla 5.15: Valores de las constantes de acoplamiento experimentales medidas por Kishi para los C-disacáridos $\beta(1\rightarrow4)$.

		MM3*	NAMFIS
C-celobiosa	<i>exo-anti</i>	81	53
	<i>exo-syn(-)</i>	7	3
	<i>exo-syn(+)</i>	10	25
	<i>non exo- syn(-)</i>	1	0
	<i>anti-syn(+)</i>	1	19
C2'-epi-C-celobiosa	<i>eExo-anti</i>	68	40
	<i>exo-syn(-)</i>	12	11
	<i>exo-syn(+)</i>	10	23
	<i>non exo-syn(-)</i>	9	4
	<i>non exo-anti</i>	1	22
C5-deshidroximetil-C-celobiosa	<i>exo-anti</i>	13	30
	<i>exo-syn(-)</i>	6	5
	<i>exo-syn(+)</i>	77	47
	<i>non exo- syn(-)</i>	1	0
	<i>anti-syn(+)</i>	3	18

Tabla 5.16: Distribución de población por cálculos de mecánica molecular (MM3* y $\epsilon = 80$) y la obtenida mediante el método JAMFIS para los C-disacáridos $\beta(1\rightarrow4)$.

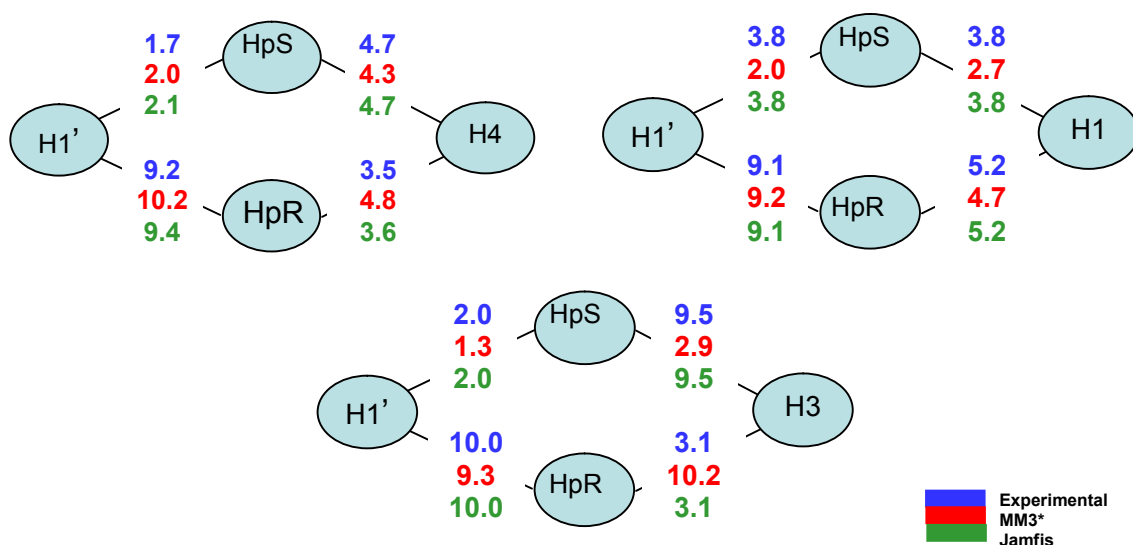


Figura 5.29: Constantes de acoplamiento experimentales, teóricas y determinadas por JAMFIS para la C-celobiosa (arriba izquierda), C2'-epi-C-celobiosa (arriba derecha) y C5-deshidroximetil-C-celobiosa (abajo).

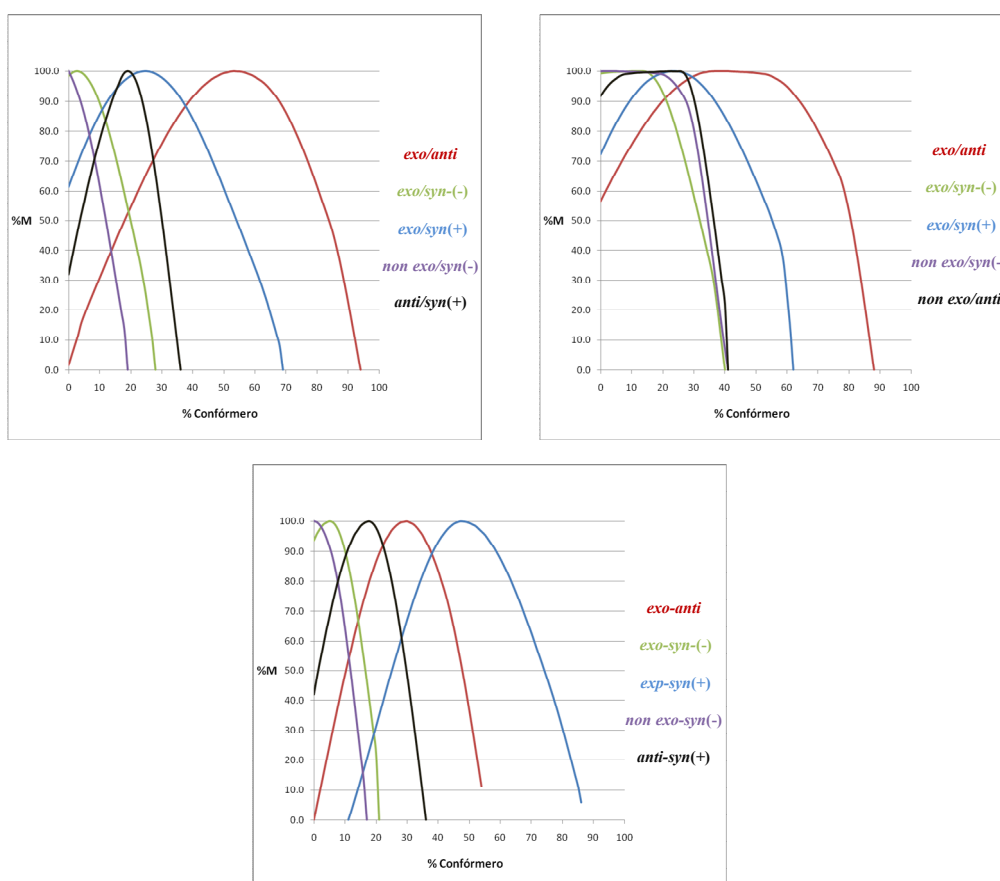


Figura 5.30: Representación del valor de M en función de la población de cada conformero para los tres C-disacáridos con unión $\beta(1\rightarrow4)$: C-celobiosa(arriba izquierda), C2'-epi-C-celobiosa (arriba derecha) y C5-deshidroximetil-C-celobiosa (abajo).

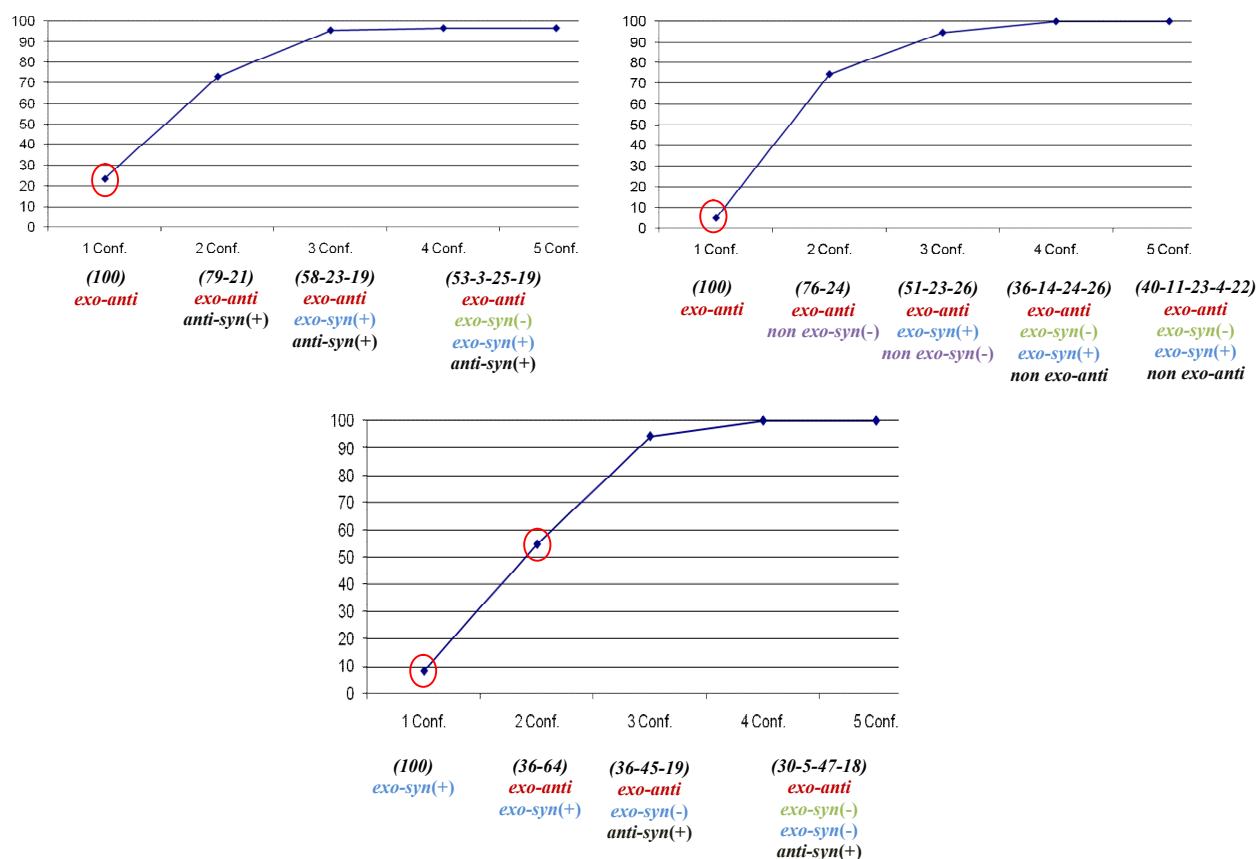


Figura 5.31: Representación del valor de M frente al número de confórmers para los tres C-disacáridos con unión $\beta(1\rightarrow4)$: C-celobiosa(arriba izquierda), C2'-epi-C-celobiosa (arriba derecha) y C5-deshidroximetil-C-celobiosa (abajo).

C-celobiosa: La distribución de población que proporciona el mejor ajuste de JAMFIS es 53:3:25:19:0 correspondiente a los confórmers *exo-anti*, *exo-syn(-)*, *exo-syn(+)*, *anti-syn(+)* y *non exo-syn(+)*, respectivamente, lo que sugiere que los confórmers, *exo-syn(-)* y *non exo-syn(+)* apenas están poblados en disolución. El análisis de todas las soluciones posibles indica que los confórmers *exo-anti*, *exo-syn(+)*, *anti-syn(+)* presentan curvas de tipo 3, por lo que los tres deben estar poblados en disolución para obtener un buen ajuste, siendo el *exo-anti* el conformero mayoritario, como en la C-lactosa. (Figura 5.30). Los otros dos confórmers presentan curvas de tipo 1, por lo que si existen en disolución, su población sería muy pequeña. Como puede observarse el valor de M mejora de forma significativa cuando se tienen en cuenta tres confórmers con respecto a utilizarse sólo uno o dos de ellos, mientras que no se ve afectado si se incluye un cuarto o quinto conformero. (Figura 5.31). Estos resultados son muy similares a los obtenidos para la C-lactosa, como podría esperarse, ya que ambos C-disacáridos son epímeros en posición C4', lo

que no afecta prácticamente al equilibrio conformacional en torno a las uniones glicosídicas.

C2'-epi-C-celobiosa: La mejor solución de JAMFIS en este caso es 40:11:23:4:22 correspondiente a *exo-anti:exo-syn(-):exo-syn(+):non exo-syn(-):non exo-anti* respectivamente, que es parecida a la distribución obtenida para la C-celobiosa salvo por la presencia del conformero *non exo/anti* que sustituye al conformero *anti-syn(+)*. La gráfica que representa todas las soluciones es también similar a la de la C-celobiosa (salvo por el cambio *non exo/anti* en lugar de la *anti-syn(+)*) con dos curvas de tipo 3, la correspondiente al conformero *exo-anti*, que es el mayoritario, y la del conformero *exo-syn(+)*, por lo que ambas familias conformacionales estarían pobladas. Las curvas de los otros tres conformeros son de tipo 1, pero muy anchas; es decir están poco definidas. No obstante, el análisis del valor de M frente al número de conformeros indica que serían necesarios al menos tres conformeros para conseguir un buen ajuste.

C5-deshidroximetil-C-celobiosa: La mejor solución de JAMFIS es 30:5:47:18:0 correspondiente a *exo-anti:exo-syn(-):exo-syn(+):anti-syn(+):non exo-syn(-)* respectivamente. En este caso las curvas están bien definidas, con tres curvas de tipo 3, correspondiente a los conformeros *exo-syn(+)*, *exo-anti* y *anti-syn(+)* que estarían poblados en disolución, aunque en este caso el conformero *exo-syn(+)* sería el mayoritario, en lugar del *exo-anti*. Los otros dos conformeros presentan curvas de tipo 1. Es decir, no estarían poblados o su población sería pequeña. De nuevo, la representación del valor de M frente al número de conformeros indica que son necesarios tres conformeros para obtener constantes de acoplamiento que sean consistentes con los valores experimentales, ya que con sólo dos de ellos no se encuentra una solución compatible con las constantes de acoplamiento experimentales.

La presencia de tres familias conformacionales en disolución para los tres C-disacáridos, pone de manifiesto la flexibilidad de la unión β -C-(1 \rightarrow 4), tanto en torno al enlace glicosídico (orientaciones *exo* y *anti- ϕ*) como en torno al aglicónico (orientaciones *syn* y *anti- ψ*) Este hecho contrasta con la predominancia del conformero *exo/syn* en los O-disacáridos. Como en el caso de la unión α -C-(1 \rightarrow 4),

el cambio de orientación del hidroxilo en C2' de ecuatorial a axial provoca la aparición de un porcentaje apreciable de conformaciones *non exo*-anoméricas, que apenas existen en los compuestos naturales, mientras que la eliminación del hidroximetilo en C5 provoca un cambio conformacional significativo en torno al enlace aglicónico. Así, mientras que en la C-celobiosa y la C2'-epi-C-celobiosa el conformero *exo-anti* es el mayoritario, en el derivado sin el hidroximetilo el conformero *exo-syn* es el más abundante, aunque sin llegar a ser tan predominante como en la celobiosa natural.

5. Conclusiones:

En una serie de artículos publicados a finales de los 80 y principios de los 90, Kishi y colaboradores analizaron el comportamiento conformacional de un gran número de C-disacáridos mediante una aproximación cualitativa basada en constantes de acoplamiento protón-protón. En este capítulo hemos abordado el análisis conformacional de algunos de estos C-disacáridos, utilizando, como él, constantes de acoplamiento escalares vecinales protón-protón como única información experimental. No obstante, nuestra aproximación ha permitido llegar a obtener porcentajes cuantitativos de las poblaciones presentes en los distintos equilibrios conformacionales. Esto se debe a la sinergia existente en nuestra aproximación entre los datos experimentales y los obtenidos mediante métodos teóricos. En primer lugar, las geometrías de los confórmeros a tener en cuenta se determinaron mediante cálculos de mecánica molecular aunque prescindiendo de sus energías relativas, ya que éstas son menos fiables y muy dependientes del campo de fuerzas utilizado. En segundo lugar, se calcularon las constantes de acoplamiento promedio de todas las posibles combinaciones de confórmeros para encontrar que distribución de población (método 1), que rango de población de cada confórmero (método 2) y que número de confórmeros (método 3) permitía el mejor ajuste con los valores experimentales. Antes de aplicar esta metodología para los C-disacáridos analizados por Kishi, el método se validó mediante su aplicación a tres C-disacáridos estudiados exhaustivamente en nuestro grupo de investigación. Para estas moléculas, el protocolo JAMFIS proporcionó conclusiones idénticas o similares a las obtenidas mediante la combinación de cálculos de mecánica molecular y NOEs.

A modo de resumen, los resultados obtenidos para cada C-disacárido se recogen en la **Tabla 5.20**, indicándose la conformación mayoritaria en cada caso y los confórmeros poblados en disolución a partir del análisis llevado a cabo con el método 2. En la tabla también se incluyen los resultados de Kishi para facilitar la comparación entre ambos estudios.

Tipo de unión		Kishi	Conformación Mayoritaria	Otras Conformaciones
Glc-Glc	$\alpha, \alpha-(1 \rightarrow 1)$	<i>exo/exo</i>	<i>exo/exo</i>	
	$\alpha, \beta-(1 \rightarrow 1)$	<i>exo/exo</i>	<i>exo/exo</i>	
	$\beta, \beta-(1 \rightarrow 1)$	<i>exo/mezcla</i>	<i>exo/anti</i>	<i>exo/exo</i> <i>non exo/exo</i>
Man-Man	$\alpha-(1 \rightarrow 2)$	<i>No analizado</i>	<i>exo/syn(+)</i>	<i>exo/syn(-)</i> <i>non exo/syn(+)</i>
Gal-Gal	$\alpha-(1 \rightarrow 3)$	<i>exo-syn(-)</i>	<i>exo-syn(-)</i>	
Gal-Glc		<i>Exo/syn distorsionada</i>	<i>exo/anti</i>	<i>exo/syn(-)</i> <i>non exo/syn(+)</i>
Gal-Man		<i>exo/mezcla</i>	<i>exo/syn(+)</i>	<i>exo/syn(-)</i>
Gal-Glc	$\beta-(1 \rightarrow 3)$	<i>No analizado</i>	<i>exo/anti</i>	<i>exo/syn(+)</i> <i>anti/syn(+)</i>
Glc-Glc Maltosa	$\alpha-(1 \rightarrow 4)$	<i>exo/mezcla</i>	<i>exo/anti</i>	<i>exo/syn(-)</i> <i>non exo/syn(+)</i>
Man-Glc		<i>exo/mezcla</i>	<i>exo/anti</i>	<i>exo/syn(-)</i>
3-deoximaltosa		<i>exo/syn(-)</i>	<i>exo/syn(-)</i>	
Gal-Glc Lactosa	$\beta-(1 \rightarrow 4)$	<i>exo/mezcla</i>	<i>exo/anti</i>	<i>exo/syn(+)</i>
Glc-Glc Celobiosa		<i>exo/mezcla</i>	<i>exo/anti</i>	<i>exo/syn(+)</i> <i>anti/syn(+)</i>
Man-Glc		<i>exo/mezcla</i>	<i>exo/anti</i>	<i>exo/syn(+)</i>
5-deshidroximetil- celobiosa		<i>exo-anti</i>	<i>exo/syn(+)</i>	<i>anti/syn(+)</i> <i>exo/anti</i>

Tabla 5.20: Comparativa de las conclusiones obtenidas en los estudios de Kishi frente a las obtenidas en este trabajo. Los conformeros presentes en disolución son aquellos que presentaban curvas de tipo 3 en el análisis de JAMFIS siguiendo el método 2.

Si se comparan los resultados de Kishi con los nuestros, nos encontramos con dos casos límites:

CASO 1: Existen unos pocos C-disacáridos para los que ambos métodos predicen que la conformación *exo-syn* (denominada *exo-exo* en enlaces $1 \rightarrow 1$) es muy predominante (>90%), por lo que su comportamiento conformacional es similar al de

los correspondientes *O*-disacáridos. Este es el caso de las dos *C*-trehalosas simétricas, del *C*-disacárido α -*C*-Gal-(1→3)-Gal y de la 3-deoxi-*C*-maltosa. En este caso, el análisis de JAMFIS permite confirmar las conclusiones obtenidas previamente por el grupo de Y. Kishi.

CASO 2: Existen *C*-disacáridos que, según el grupo de Harvard, adoptan una conformación *exo*-anomérica en torno al enlace glicosídico, mientras que en torno al enlace aglicónico existen como una mezcla indefinida de conformeros alternados, o que adoptan una conformación distorsionada con respecto a la alternada. En estos casos el análisis de JAMFIS ha permitido caracterizar el comportamiento conformacional de los *C*-disacáridos con mucho más detalle y en muchos casos se ha comprobado la existencia de flexibilidad tanto alrededor del enlace glicosídico (presencia de conformeros *anti*- ϕ y *non exo*) como del aglicónico (presencia de conformeros *syn*(+), *syn*(-) y *anti*- ψ). La flexibilidad de estas moléculas se manifiesta alrededor de ambos enlaces, glicosídico y aglicónico.

Por tanto, la conclusión general del grupo de Y. Kishi de que los *C*-disacáridos presentan las mismas propiedades conformacionales que los *O*-disacáridos sólo es correcta para un pequeño número de *C*-disacáridos (caso A), que presentan un comportamiento prácticamente mono-conformacional en disolución. El resto de los *C*-disacáridos son mucho más flexibles que los correspondientes *O*-disacáridos con la existencia de varias familias conformacionales en disolución. De hecho, existen *C*-disacáridos que presentan conformaciones mayoritarias de tipo *anti*- ψ en el enlace aglicónico y porcentajes en torno al 20% de conformeros *anti*- ϕ y *non exo* en el glicosídico. Estas geometrías son muy raras en los glicósidos naturales.

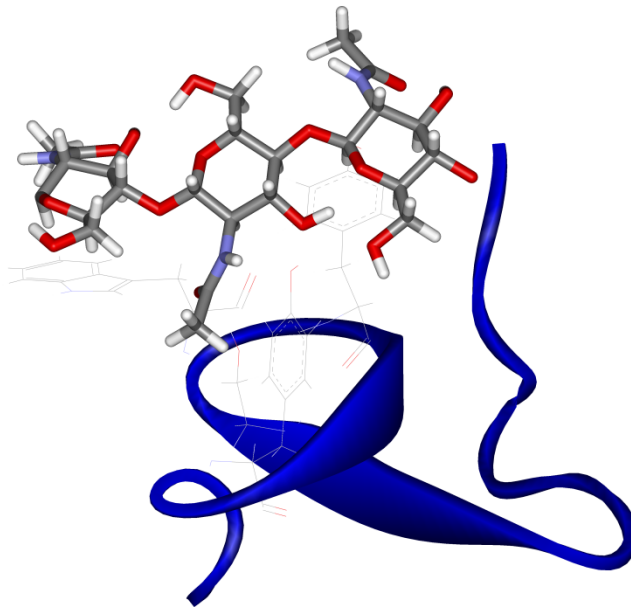
Este resultado pone de manifiesto la importancia del efecto *exo*-anomérico en definir las propiedades conformacionales de los disacáridos, ya que la ausencia de esta estabilización estéreo-electrónica produce cambios conformacionales y aumento de la flexibilidad en torno al enlace glicosídico. La mayor flexibilidad alrededor del enlace aglicónico debe ser debida a la mayor longitud del enlace C-C frente al C-O, que separa ligeramente los dos anillos de piranosa que flanquean el enlace glicosídico.

6. Bibliografía:

- (1) Goekjian, P. G.; Wei, A.; Kishi, Y. In *Carbohydrate-Based Drug Discovery*; Wong, C.-H., Ed.; Wiley-VCH 2003.
- (2) Wu, T. C.; Goekjian, P. G.; Kishi, Y. *J. Org. Chem.* **1987**, *52*, 4819-4823.
- (3) Goekjian, P. G.; Wu, T. C.; Kang, H. Y.; Kishi, Y. *J. Org. Chem.* **1987**, *52*, 4823-4825.
- (4) Babirad, S. A.; Wang, Y.; Goekjian, P. G.; Kishi, Y. *J. Org. Chem.* **1987**, *52*, 4825-4827.
- (5) Wang, Y.; Goekjian, P. G.; Ryckman, D. M.; Kishi, Y. *J. Org. Chem.* **1988**, *53*, 4151-4153.
- (6) Miller, W. H.; Ryckman, D. M.; Goekjian, P. G.; Wang, Y.; Kishi, Y. *J. Org. Chem.* **1988**, *53*, 5580-5582.
- (7) Goekjian, P. G.; Wu, T. C.; Kishi, Y. *J. Org. Chem.* **1991**, *56*, 6412-6422.
- (8) Goekjian, P. G.; Wu, T. C.; Kang, H. Y.; Kishi, Y. *J. Org. Chem.* **1991**, *56*, 6422-6434.
- (9) Wang, Y.; Babirad, S. A.; Kishi, Y. *J. Org. Chem.* **1992**, *57*, 468-481.
- (10) Wang, Y.; Goekjian, P. G.; Ryckman, D. M.; Miller, W. H.; Babirad, S. A.; Kishi, Y. *J. Org. Chem.* **1992**, *57*, 482-489.
- (11) Haneda, T.; Goekjian, P. G.; Kim, S. H.; Kishi, Y. *J. Org. Chem.* **1992**, *57*, 490-498.
- (12) O'Leary, D. J.; Kishi, Y. *J. Org. Chem.* **1993**, *58*, 304-306.
- (13) Wei, A.; Kishi, Y. *J. Org. Chem.* **1994**, *59*, 88-96.
- (14) O'Leary, D. J.; Kishi, Y. *J. Org. Chem.* **1994**, *59*, 6629-6636.
- (15) Wei, A.; Haudrechy, A.; Audin, C.; Jun, H.-S.; Haudrechy-Bretel, N.; Kishi, Y. *J. Org. Chem.* **1995**, *60*, 2160-2169.
- (16) Ravishankar, R.; Surolia, A.; Vijayan, M.; Lim, S.; Kishi, Y. *J. Am. Chem. Soc.* **1998**, *120*, 11297-11303.
- (17) Jimenez-Barbero, J.; Espinosa, J. F.; Asensio, J. L.; Canada, F. J.; Poveda, A. *Adv. Carbohydr. Chem. Biochem.* **2000**, *56*, 235-84.
- (18) Weimar, T.; Woods, R. J. In *NMR Spectroscopy of Glycoconjugates*; Jiménez-Barbero, J., Peters, T., Eds.; Wiley-VCH, , 2003; pp 111-144.
- (19) Cicero, D. O.; Barbato, G.; Bazzo, R. *J. Am. Chem. Soc.* **1995**, *117*, 1027-1033.
- (20) Haasnoot, C. A. G.; de Leeuw, F. A. A. M.; Altona, C. *Tetrahedron* **1980**, *36*, 2783-2792.
- (21) Espinosa, J.-F.; Martín-Pastor, M.; Asensio, J. L.; Dietrich, H.; Martín-Lomas, M.; Schmidt, R. R.; Jiménez-Barbero, J. *Tetrahedron Lett.* **1995**, *36*, 6329-6332.
- (22) Espinosa, J.-F.; Cañada, F. J.; Asensio, J. L.; Martín-Pastor, M.; Dietrich, H.; Martín-Lomas, M.; Schmidt, R. R.; Jiménez-Barbero, J. *J. Am. Chem. Soc.* **1996**, *118*, 10862-10871.
- (23) Espinosa, J.-F.; Bruix, M.; Jarreton, O.; Skrydstrup, T.; Beau, J.-M.; Jiménez-Barbero, J. *Chem. Eur. J.* **1999**, *5*, 442-448.
- (24) Rubinstenn, G.; Sinay, P.; Berthault, P. *J. Phys. Chem. A* **1997**, *101*, 2536-2540.
- (25) Mikros, E.; Labrinidis, G.; Pérez, S. *J. Carbohydr. Chem.* **2003**, *22*, 407 - 421.
- (26) Ferritto, R.; Vogel, P. *Tetrahedron: Asymmetry* **1994**, *5*, 2077-2092.
- (27) Mikros, E.; Labrinidis, G.; Pérez, S. *J. Carbohydr. Chem.* **2000**, *19*, 1319-1349.

Capítulo 6

Dominios de Heveína: Mutantes del Ac-AMP2



1. Introducción y antecedentes:

Como se comentó anteriormente, las lectinas son proteínas que unen carbohidratos con gran selectividad, reversibilidad y eficacia¹. Estas proteínas, en muchos casos glicosiladas, y que no presentan actividad enzimática se encuentran tanto en el mundo animal como vegetal, así como en microorganismos como bacterias, protozoarios y virus (<http://www.cermav.cnrs.fr/lectines/>).

La primera lectina que se caracterizó, mediante difracción de Rayos X, fue la Concanavalina A (Con A), de origen vegetal y obtenida a partir de *Canavalia Ensiforme*. Las lectinas de origen vegetal han sido utilizadas ampliamente como herramientas en áreas de investigación biomédica y biológica². Más concretamente, las lectinas que unen quitina han sido relacionadas con la defensa de plantas frente a insectos y hongos³.

Muchas de estas lectinas incluyen un motivo estructural común denominado dominio de heveína⁴, que contiene alrededor de 40 aminoácidos y está caracterizado por la presencia, en posiciones conservadas, de 6 u 8 cisteínas, que forman 3 ó 4 puentes disulfuro. Este nombre procede de la lectina heveína, que es un polipéptido de 43 aminoácidos, que se extrajo por primera vez de la goma del látex de la *Hevea brasiliensis*⁵. Independientemente de sus propiedades de defensa en la planta de origen, se ha demostrado que la propia heveína y otras proteínas homólogas presentes en otras plantas están relacionadas con la alergia al látex y con el denominado síndrome cruzado látex-frutas⁶. Los dominios de heveína pueden encontrarse aislados, como en la propia heveína y en la pseudoheveína, o formando parte de proteínas más grandes e incluso agrupados covalentemente en conjuntos de dos o más dominios. Así, por ejemplo, la lectina de la ortiga (UDA, *Urtica dioica agglutinin*) y la aglutinina del germen de trigo (WGA, *Wheat germen agglutinin*), muestran dos y cuatro dominios de heveína, respectivamente. La actividad aglutinante de la WGA fue una de las primeras propiedades biológicas demostrada de modo inequívoco y su estructura se caracterizó mediante cristalografía de Rayos X muy tempranamente⁷. Las proteínas más pequeñas de esta familia se han aislado de las semillas de *Amarantus caudatus* y muestran propiedades antifúngicas y antimicrobianas. Estos péptidos, denominados Ac-AMP1 y Ac-AMP2 (*Amarantus*

caudatus-Antimicrobial peptide), poseen sólo 3 puentes disulfuro y tienen 29 y 30 aminoácidos, respectivamente.

Basándose en los estudios estructurales, por homología y de reconocimiento molecular, realizados sobre estos dominios, que incluye los trabajos desarrollados por nuestro grupo⁸⁻¹³, ha sido posible definir aquellos aminoácidos que son esenciales para el reconocimiento de oligómeros de *N*-acetilglucosamina. Éstos son, en numeración de heveína, la Ser19, los residuos aromáticos 21 y 23, dos triptófanos en el caso de la heveína, y la Tyr30. Estos estudios permitieron determinar la existencia de distintos tipos de fuerzas implicadas en el proceso de reconocimiento: enlaces de hidrógeno, fuerzas electrostáticas, de van der Waals e interacciones de apilamiento CH- π ¹⁴⁻¹⁵.

1. Las energías de enlace asociadas a los enlaces de hidrógeno están en el rango de 12 a 30 kJ/mol, por lo que se trata de un enlace relativamente débil. Sin embargo, entre los enlaces débiles, son considerados relativamente fuertes, ya que pueden llegar a suponer el 10% de la energía contenida en los enlaces covalentes. Por lo tanto, la presencia de numerosos grupos hidroxilo en los carbohidratos claramente facilita la participación de enlaces intermoleculares de hidrógeno con grupos donadores y aceptores de puentes de hidrógeno de la proteína¹⁶.
2. Por otro lado, aunque los azúcares sean marcadamente hidrófilos, las orientaciones específicas de sus grupos no polares C-H pueden crear interacciones hidrófobas, que pueden interactuar recíprocamente con una cavidad, también hidrófoba, del sitio de unión del receptor. Dentro de las interacciones no polares se encuentran las fuerzas de van der Waals, que son fuerzas no específicas atractivas entre dos átomos o grupos de átomos. Estas fuerzas son más favorables si los átomos se encuentran a 0,2-0,3 nm de distancia mutua, dependiendo de su naturaleza. La fuerza proviene de la atracción transitoria entre el núcleo de un átomo y la nube de electrones del átomo vecino. La energía de enlace de las fuerzas de van der Waals es de aproximadamente 4 kJ/mol. Es decir, estas fuerzas son más débiles y menos específicas que los enlaces de hidrógeno, aunque la suma de muchas interacciones de van der Waals puede apoyar, considerablemente, la formación de enlaces de hidrógeno entre el

carbohidrato y las cadenas laterales de la proteína, proporcionando una mayor estabilización del complejo¹⁷⁻¹⁸.

3. Adicionalmente, otras interacciones relevantes para el reconocimiento molecular de carbohidratos por proteínas, y que tienen componentes polares y no polares, son las interacciones entre los grupos C-H de los residuos del carbohidrato con el sistema de electrones π de residuos de aminoácidos aromáticos (triptófano, fenilalanina y tirosina). Estos residuos se localizan frecuentemente en los sitios de unión con carbohidratos. Estas interacciones de apilamiento CH- π son determinantes de la conformación y de la afinidad de los carbohidratos como ligandos, por lo que permiten la formación y modulan la selectividad y estabilidad de los complejos proteína-carbohidrato¹⁹⁻²².

Así, la Ser19 y la Tyr30 participan en la interacción mediante enlaces de hidrógeno entre sus OHs y el azúcar, mientras que existen interacciones de tipo CH- π entre los anillos de piranosa y los anillos aromáticos²³ de las posiciones 21 y 23, y entre el grupo metilo de un grupo acetamido del azúcar y el anillo aromático de la Tyr30 (Figura 6.1).

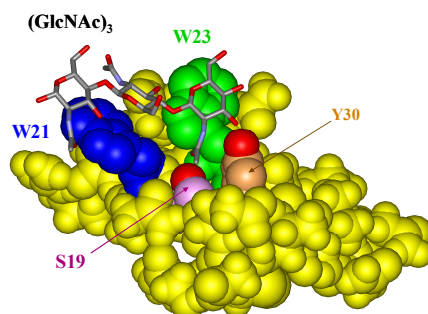
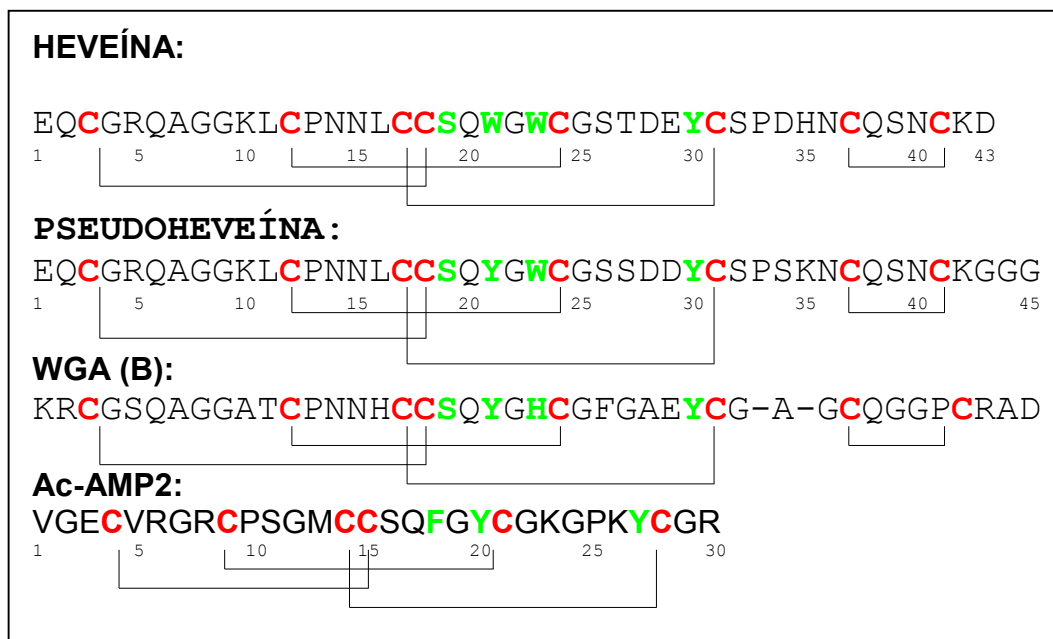


Figura 6.1. Interacciones de apilamiento CH- π y enlaces de hidrógeno entre la heveína truncada de 32 aa y la tri-N-acetilquitotriosa⁴⁹.

Tanto mediante cristalografía de Rayos X, como por RMN, se ha determinado la estructura tridimensional de distintos dominios de heveína, en estado libre o asociado a oligosacáridos. Por ejemplo, se han descrito los casos de la WGA⁷ y de su dominio B aislado, de la UDA libre²⁴ y asociada a quitotriosa²⁵, de la heveína de látex libre²⁶, asociada a quitobiosa⁸, quitotriosa²⁷ y quitopentosa¹¹, de una heveína truncada de 32 aminoácidos²⁷, de la pseudoheveína¹³, del PL-D2 (aislado de la raíz

de la fitolaca o hierba carmín, *Pokeweed*)²⁸, del péptido antifúngico EAFP2 aislado del árbol de la gutapercha, *Eucommia ulmoides*²⁹, o del péptido antifúngico Ac-AMP2³⁰ (Esquema 1.1)



Esquema 6.1: Secuencias de distintas proteínas con dominio Heveína. En verde se marcan los aminoácidos implicados en la interacción carbohidrato-proteína.

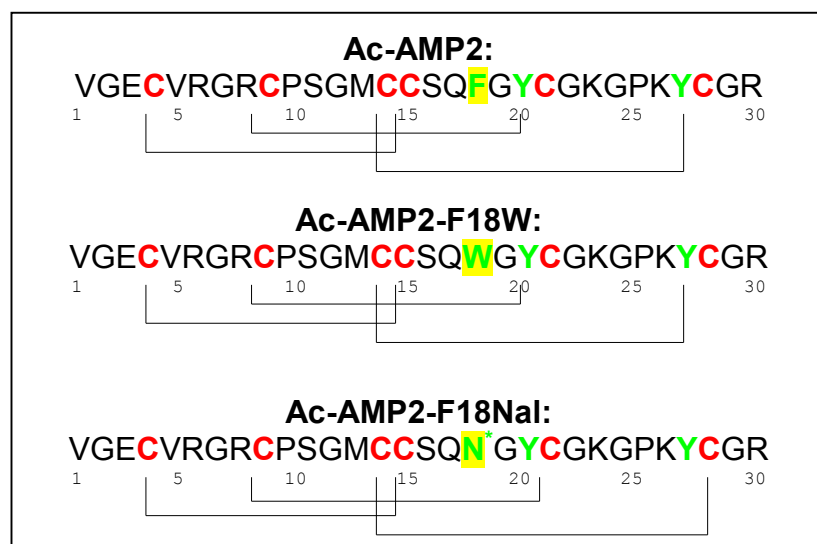
Por otro lado, para entender, de la manera más rigurosa posible, los factores involucrados en el proceso de reconocimiento molecular de carbohidratos por proteínas, es fundamental conocer los parámetros termodinámicos del proceso de asociación. En este sentido, se ha demostrado que, de modo general, existe un fenómeno de compensación entalpía-entropía, que se ha atribuido a la pérdida de grados de libertad conformacional tanto de la proteína como del carbohidrato al formarse el complejo³¹.

Uno de los intereses y objetivos fundamentales de nuestro grupo de investigación es el investigar y comprender el origen de la especificidad y estabilidad de los complejos proteína-carbohidrato, profundizando en el conocimiento de la importancia que juegan las interacciones de apilamiento carbohidrato-aromático y los enlaces de hidrógeno en el proceso de reconocimiento molecular. Puesto que los dominios de heveína son relativamente pequeños y manejables, ha sido posible, en colaboración con otros grupos de investigación, disponer de heveínas naturales o

mutantes mediante métodos de purificación de extractos³², de biología molecular³³ o de síntesis química³⁴⁻³⁵. Además, esta metodología también ha permitido sintetizar péptidos modificados artificialmente, donde algún aminoácido del sitio de unión ha sido sustituido químicamente por otro aminoácido, natural o no natural. Esta mutación permite estudiar la importancia de la naturaleza química de los diversos residuos para la unión a azúcares. De este modo, se han conseguido sintetizar péptidos de tipo Ac-AMP2 mutados en su posición 18 (21, según la nomenclatura de la heveína), sustituyendo la fenilalanina del Ac-AMP2 por un triptófano o por aminoácidos no naturales, como la β -2-naftil-alanina o 4-fluorofenilalanina.

En este capítulo se abordará el estudio de la estructura tridimensional mediante RMN de los mutantes de Ac-AMP2-F18W y Ac-AMP2-F18Nal, sintetizados químicamente por Michiro Muraki³⁴. En estos péptidos, se ha modificado la naturaleza química del aminoácido aromático, 18 (equivalentes a 21 en heveína) implicado en el sitio de unión con el carbohidrato (**Esquema 6.2**).

Asimismo se ha estudiado la interacción, tanto de la proteína nativa como de estos dos mutantes, con *N,N',N''*-triacetilquitotriosa (GlcNAc)₃ determinando las constantes de asociación y los parámetros termodinámicos de van't Hoff correspondientes a la interacción.



Esquema 6.2: Secuencia de aminoácidos del AcAMP2 y los mutantes estudiados en esta tesis.

2. Determinación de la estructura 3D del mutante Ac-AMP2-F18W

Como se ha explicado en el capítulo 1, el análisis combinado de espectros TOCSY y NOESY proporciona la asignación completa de todas o de la mayoría de las resonancias de protón del péptido. En nuestro caso, la asignación de estos espectros se realizó basándose en la ya descrita³⁰ para la proteína nativa. Para la determinación de la estructura de este péptido, se realizaron distintos experimentos 2D a 500 y 800 MHz, tanto en estado libre como asociado al azúcar. Algunos espectros se representan en la **Figura 6.2**, **Figura 6.3**, **Figura 6.4** y **Figura 6.5**.

2.1. Análisis de la proteína libre:

Con la combinación de la información obtenida de estos espectros llegamos a la asignación completa del péptido (**Tabla 6.1**).

	HN	α H	β H	γ H	Otros
Val1		3.87	2.19	1.01	
Gly2	8.91	3.98, 4.10			
Glu3	8.69	4.28	1.88, 1.97	2.18, 2.29	
Cys4	7.75	4.42	2.75, 3.09		
Val5	8.34	4.08	1.86	0.84	
Arg6	9.30	3.84	1.79, 1.88	1.56	H δ : 3.15, H ϵ : 6.69, 7.08
Gly7	8.29	3.69, 4.08			
Arg8	7.84	4.63	1.82, 1.87	1.53, 1.61	H δ : 3.13
Cys9	8.58	5.17	2.25, 2.80		
Pro10		4.40	1.68, 2.35	1.94	H δ : 3.91, 3.42
Ser11	8.31	4.11	3.79		
Gly12	8.86	3.63, 4.19			
Met13	7.91	4.63	1.71, 1.99	2.29	
Cys14	9.2	4.57	2.36, 3.80		
Cys15	8.76	4.72	2.84		
Ser16	9.63	4.81	4.06, 4.20		
Gln17	8.99	3.94	1.47, 1.64	1.38, 1.57	H ϵ : 7.59, 7.05
Trp18	7.28	4.80	3.15, 3.62		Aromáticos: 7.26, 7.64, 10.17, 7.12,
Gly19	7.79	3.53, 4.00			
Tyr20	7.39	5.02	2.83, 3.23		Aromáticos: 6.85, 6.99
Cys21	8.80	5.48	2.77		
Gly22	8.43	1.87, 3.64			
Lys23	8.03	4.90	1.62, 1.83	1.30, 1.41	H δ : 1.63, H ϵ : 2.86
Gly24	8.33	3.94, 4.55			
Pro25		4.25	1.92, 2.26	2.03, 2.04	H δ : 3.81, 3.62
Lys26	8.79	4.03	1.81	1.13, 1.39	H δ : 1.60, H ϵ : 2.86
Tyr27	7.55	4.05	2.50, 2.90		H δ : 7.15, H ϵ : 6.71
Cys28	8.45	4.51	2.71, 3.20		
Gly29	7.98	3.95, 3.98			
Arg30	8.28	4.25	1.68, 1.85	1.58	H δ : 3.14

Tabla 6.1: Desplazamientos químicos (δ) de ¹H-RMN del Ac-AMP2-F18Nal (25° C, pH 5,6)

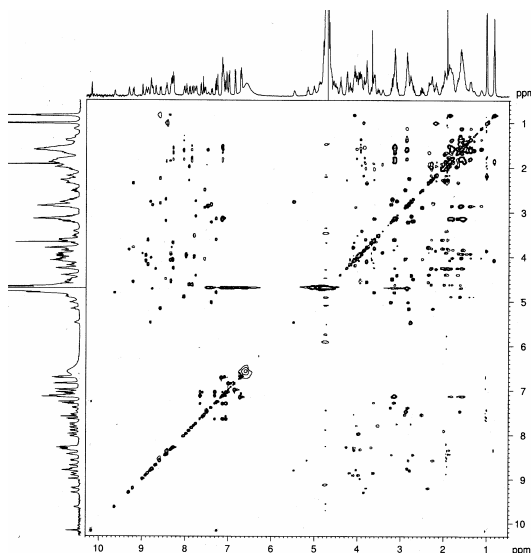


Figura 6.2: Espectro TOCSY a 500MHz del Ac-AMP2 F18W en H₂O:D₂O (90:10) y tiempo de mezcla de 60 ms libre.

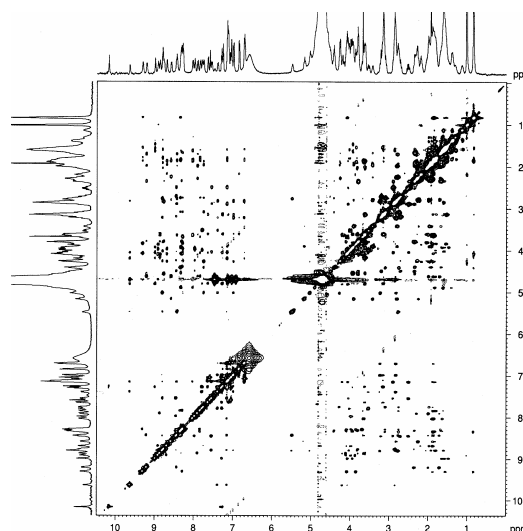


Figura 6.3: Espectro NOESY a 500 MHz del Ac-AMP2 F18W en H₂O:D₂O (90:10) y tiempo de mezcla 250 ms.

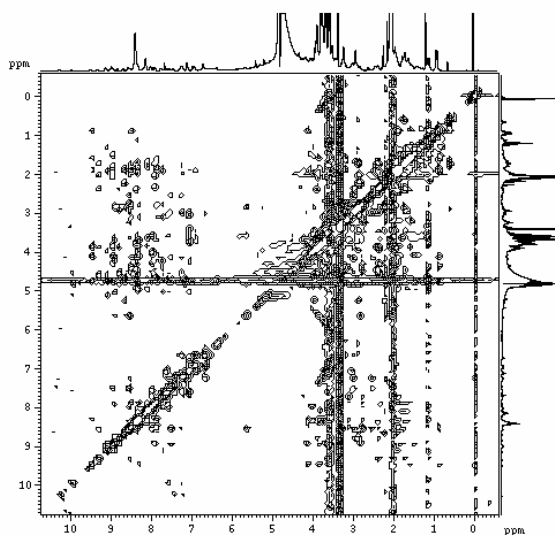


Figura 6.4: Espectro NOESY a 800 MHz del Ac-AMP2 F18W: (GlcNAc)₃ (8:1) y tiempo de mezcla de 200 ms.

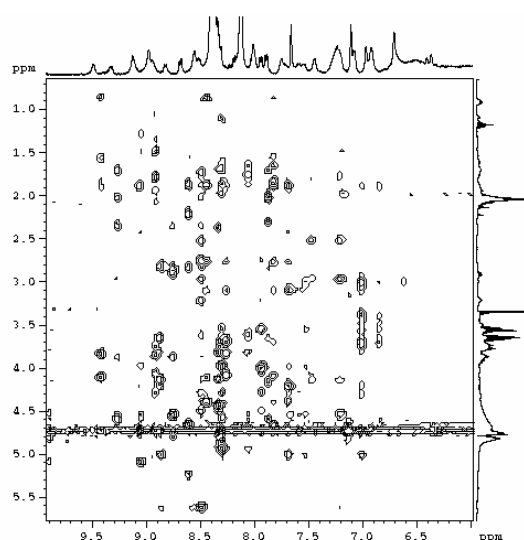
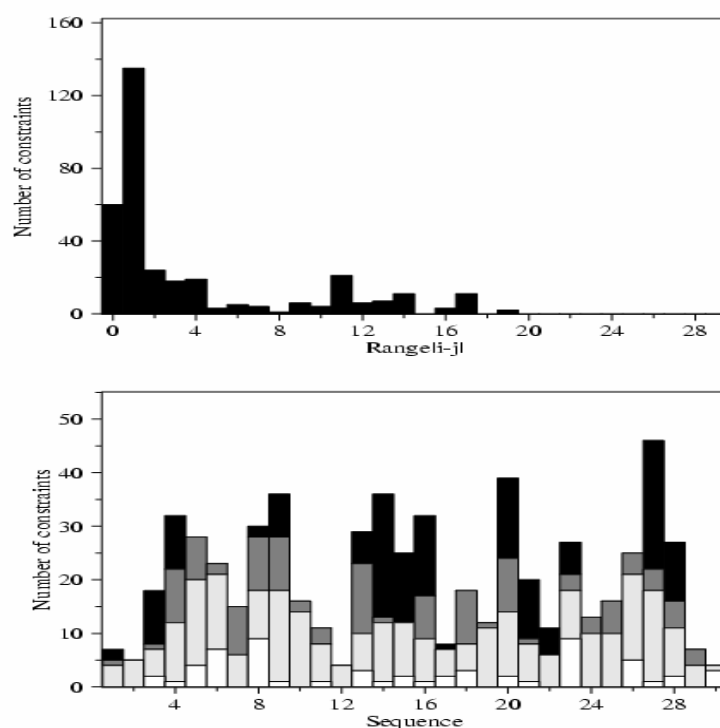


Figura 6.5: Ampliación del espectro NOESY a 800 MHz del Ac-AMP2 F18W asociado al azúcar

2.1.1. Protocolo de identificación de estructuras secundarias:

El análisis de los experimentos NOESY con distintos tiempos de mezcla y/o campos magnéticos permitió la asignación de 331 NOEs informativos, además de 67 NOEs adicionales no significativos desde un punto de vista estructural. Los NOEs clave se pueden clasificar en: 60 intra-residuos, 135 secuenciales, 59 de rango medio ($r < 5$) y 77 de rango largo ($r > 5$). Las

distribuciones de restricciones en función de su rango, y en función del tipo de residuo se representan en las **Figuras 6.6 y 6.7**, respectivamente.



Figuras 6.6 y 6.7: Distribución de las restricciones halladas por RMN. En la figura 6.6 (arriba) podemos ver el número de restricciones que hay dependiendo del tipo: secuenciales, a media distancia y a larga distancia. En la figura 6.7 se representa el tipo de restricciones que hay por aminoácido.

En un primer paso se determinó la estructura secundaria. En concreto, fue posible encontrar un número de parámetros de RMN característicos de la existencia de una lámina β anti-paralela mientras que no se encontraron indicios de existencia de hélices.

Los parámetros de RMN característicos de este elemento estructural son³⁶:

- NOEs secuenciales: $\alpha N(i, i+1)$ intensos
- NOEs a larga distancia: $\alpha \alpha (i, j)$ débiles

Además, en este tipo de estructura secundaria, existe una diferencia significativa del desplazamiento químico (δ) de los protones α con respecto a los de referencia para un ovillo estadístico ($\Delta\delta_{H\alpha}$ positivos para una serie de residuos consecutivos).

De acuerdo con los datos de NOE y de desplazamiento químico, los aminoácidos implicados en una lámina β van desde la Met13 a la Lys23. Por

ejemplo, en la **Tabla 6.2**, se aprecia que los $H\alpha$ de Ser16 y Tyr20 del AcAMP2 F18W aparecen a 4.81 y 5.02 ppm, en vez de a sus δ básicos, 4.50 y 4.60 ppm, respectivamente. Además, existen NOEs característicos de una lámina β : $H\alpha_{15}$ -HN16, $H\alpha_{21}$ -HN22 (tipo αH_i - NH_{i+1}), $H\alpha_{15}$ - $H\alpha_{21}$ (tipo αH_i - αH_j) y HN14- $H\alpha_{23}$, HN15- $H\alpha_{21}$ (tipo NH_i - αH_j).

Aminoácido	Valor Experimental	Valor teórico
Met 13	4.63	4.52
Ser 16	4.81	4.50
Gly 19	3.53, 4.00	3.97
Tyr 20	5.02	4.60
Lys23	4.90	4.36

Tabla 6.2: Valores experimentales esperados para un ovillo estadístico de algunos $H\alpha$ del AcAMP2 F18W implicados en la lámina β .

2.1.2. Protocolo de generación de estructuras 3D

Como ya se ha mencionado, el análisis de los diferentes espectros NOESY permitió la asignación de 331 NOEs en el mutante F18W. Estos NOEs se tradujeron a restricciones de distancia a partir de la estimación de la intensidad de los picos de cruce, usando la aproximación de espines aislados. La intensidad de los picos se estimó utilizando el programa XEASY³⁷. Para ello se utilizó un criterio cualitativo: los picos muy intensos se catalogaron como correspondientes a pares de protones que están situados a una distancia menor de 2.5 Å, los picos de intensidad media, situados a menos 3.5 Å y los picos débiles, situados a menos de 5 Å¹¹.

Las estructuras 3D se generaron usando el programa DYANA³⁸, a partir de las restricciones de NOE antes mencionadas. Además de estas restricciones de NOE, también se incluyeron los puentes disulfuros de forma explícita en el cálculo. Para ellos, se restringieron las distancias S-S del siguiente modo: 2.0 Å < R < 2.1 Å y las C α -S como: 3.0 Å < R < 3.1 Å.

EL programa DYANA emplea un protocolo de enfriamiento simulado (simulated annealing (SA)), combinado con un cálculo en el espacio multidimensional de los ángulos de torsión del péptido. Así, empleando distintas estructuras iniciales, se generaron conjuntos de 400 y 800 estructuras,

de las que, en cada caso, se seleccionaron las 25 mejores. El criterio de selección se basó en su función diana, es decir, aquellas estructuras que proporcionan menos violaciones de los datos experimentales.

En la **Tabla 6.3** se detallan los datos obtenidos a partir de DYANA mientras que en la **Figura 6.8** se representa una superposición de los distintos conjuntos de estructuras obtenidas. Puede observarse que la estructura del polipéptido está bien definida tanto en la cadena principal (RMSD < 0.75 Å), como en el conjunto de los átomos (RMSD < 1.6 Å). La orientación de los aminoácidos supuestamente implicados (S16, W18, Y20 y Y27) en el reconocimiento de azúcares también está bien definida.

Nº Estructuras	Rango superposición	RMSD (Å) Cadena principal	RMSD (Å) Átomos pesados	Función dyana
25	4-28	0.73±0.23	1.53±0.30	0.43 (0.33-0.50)
	13-23(Lámina β)	0.58±0.16	1.57±0.35	

Tabla 6.3: Datos estadísticos y bondad de las estructuras obtenidas mediante cálculos DYANA para el mutante Ac-AMP2-F18W (RMSD (Å) es la desviación cuadrática media).

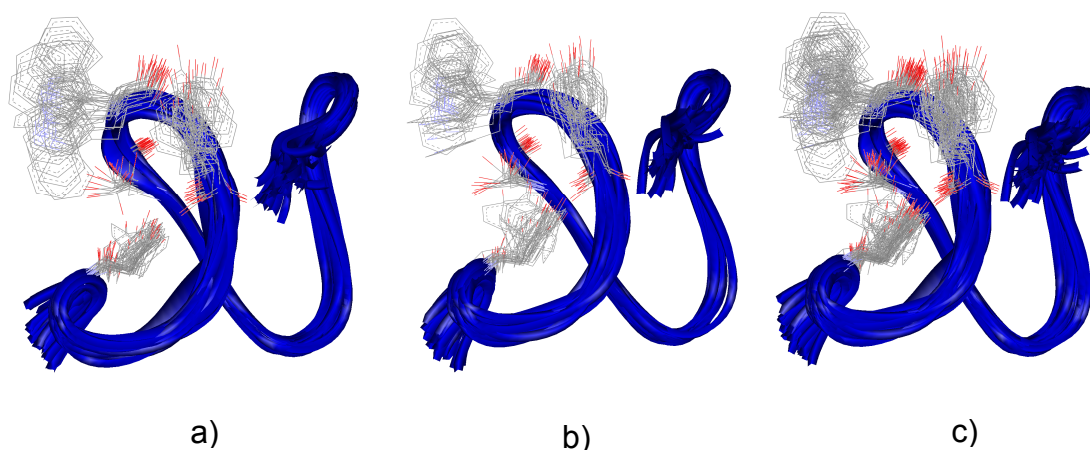


Figura 6.8: Representación esquemática de las estructuras (esqueleto y residuos implicados en el reconocimiento de quitooligosacáridos) del mutante Ac-AMP2-F18W, obtenidas mediante DYANA:

- a) Las 25 mejores estructuras del cálculo A
- b) Las 25 mejores estructuras del cálculo B
- c) Superposición de las 50 estructuras seleccionadas.

El método DYANA no optimiza los contactos y enlaces del polipéptido. Por tanto, y como paso final, cada una de estas 50 estructuras seleccionadas se optimizó mediante un protocolo de dinámica molecular restringida (RMD)³⁹, utilizando el campo de fuerzas AMBER⁴⁰. El proceso de refinamiento consistió en una minimización restringida (con los NOEs), seguida de un protocolo de calentamiento y enfriamiento simulado, para finalmente terminar con otra minimización restringida. Tras el protocolo de RMD aplicado a las 50 estructuras DYANA, se escogieron las 25 mejores en función de las violaciones de los datos experimentales. (Tabla 6.4 y Figura 6.8).

Dinámica Molecular (Amber) 25 estructuras	Rango	RMSD (Å)	RMSD (Å)
	superposición	Cadena principal	Átomos pesados
	4-27	0.80±0.23	1.86±0.41
	13-23(Lámina β)	0.42±0.13	1.87±0.45

Tabla 6.4: Datos estadísticos y bondad de las estructuras obtenidas mediante cálculos AMBER para el mutante *Ac-AMP2-F18W* (RMSD (Å) es la desviación cuadrática media).

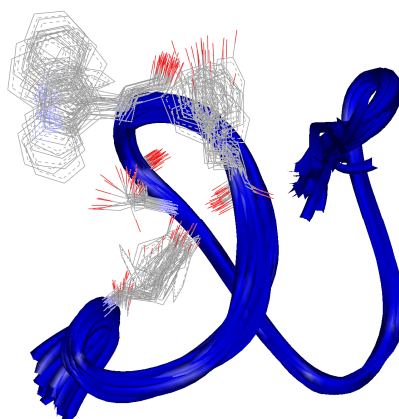


Figura 6.8: Representación esquemática de las 25 mejores estructuras (esqueleto y residuos implicados en el reconocimiento de quitooligosacáridos) del mutante *Ac-AMP2-F18W*, obtenidos mediante AMBER.

Puede observarse que los RMSDs de la lámina β disminuyen sensiblemente (pasan de 0.6 a 0.4 Å), indicando una mayor definición para este elemento de estructura secundaria que para el resto de la cadena polipeptídica. Los RMSDs para todo el esqueleto (0.8 Å) y los átomos pesados (1.9 Å) se mantienen en unos niveles más que aceptables, indicando la buena definición de las estructuras obtenidas.

Un modo usual de estimar la bondad de las estructuras obtenidas es el de representar los ángulos ϕ y ψ del péptido en el mapa de Ramachandran, que denota las áreas ϕ , ψ permitidas energéticamente para los distintos elementos estructurales que forman una cadena polipeptídica. El mapa correspondiente a las 25 estructuras proporcionadas por AMBER se muestra en la **Figura 6.9**. Los puntos representan las glicinas, y las cruces representan el resto de aminoácidos. Puede observarse, que prácticamente todos los valores ϕ , ψ se encuentran en zonas permitidas, salvo algunas glicinas (conformacionalmente más móviles y, por tanto, con posibilidad de acceso a distintas zonas del mapa) y los residuos 5, 6 y 11 de algunas estructuras, que se encuentran en la región *N*-terminal, más flexible, de la cadena del polipéptido, tal y como también se encuentra en los dominios de la heveína, la pseudoheveína y el dominio B de la aglutinina del germen del trigo¹².

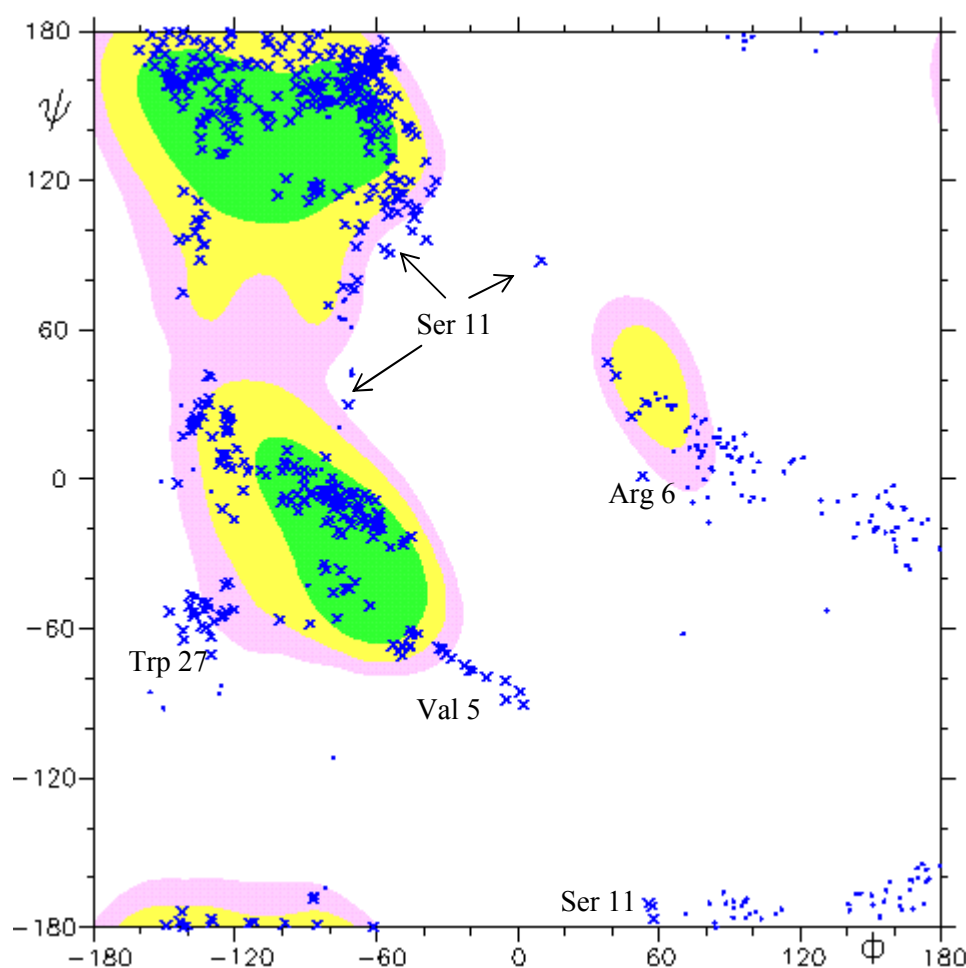


Figura 6.9: Mapa de Ramachandran de las 25 estructuras AMBER del mutante *Ac-AMP2-F18W*.

2.2. Análisis de la proteína asociada a (GlcNAc)₃

Una vez obtenida la estructura 3D del péptido en estado libre, se procedió al estudio de su interacción con *N,N',N''*-triacetilquitotriosa (GlcNAc)₃. La existencia de interacción se dedujo a partir de las variaciones de desplazamientos químicos (δ) observadas en los protones de la proteína, al añadir cantidades crecientes de azúcar (véase más adelante). La mayor variación de δ observados se produjo entre los residuos 16 y 28, de acuerdo con lo observado en otros dominios de heveínas^{10-11,13}. Por tanto, el sitio de reconocimiento de azúcares se mantiene básicamente inalterado en este mutante de AcAMP2.

Con objeto de obtener la estructura del complejo Ac-AMP2-F18W con quitotriosa, se realizaron espectros NOESY para una relación molar 4:1 (GlcNAc)₃:proteína. El análisis de los espectros NOESY puso de manifiesto que no existían cambios significativos en el número ni en la intensidad de los picos de cruce de los protones de la proteína en presencia de azúcar con respecto a los mismos en la proteína libre, hecho que implica que no existen cambios conformacionales notables en el esqueleto de la proteína al pasar del estado libre al asociado. De hecho, la existencia de preorganización de los sitios de unión de azúcares es un hecho común en lectinas⁴¹. Además la posición del sitio de reconocimiento fue corroborado por la existencia de NOEs entre la proteína y el ligando. (**Figura 6.10**)

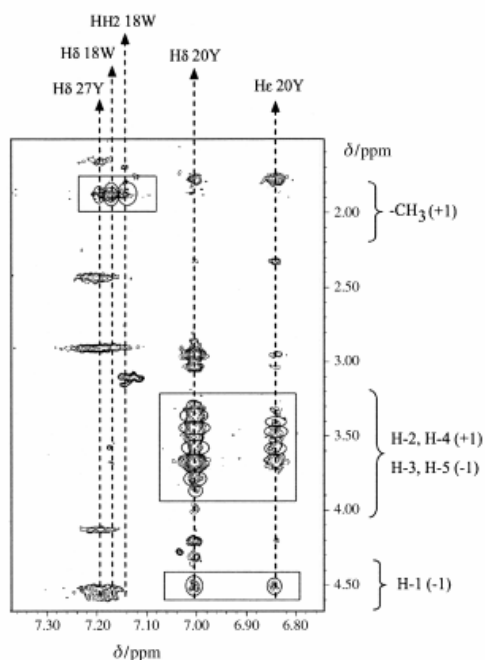


Figura 6.10: Detalle del espectro NOESY con algunos de los NOEs proteína-ligando observados.

Estudios previos de dominios de heveína con pequeños quitoooligosacáridos (por encima de tres unidades de *N*-acetil glucosamina)^{7-8,10-13,24-30} han mostrado que el lugar de reconocimiento de las lectinas tiene una extensión suficiente como para albergar un mínimo de 2 unidades de GlcNAc, colocando las dos primeras unidades en los residuos 16, 18, 20 y 27. De esta manera, en el caso del trisacárido, son posibles dos modos de unión diferentes. Como se puede apreciar en la **Figura 6.11**, estos dos modos de unión darán lugar a NOEs intermoleculares distintos.

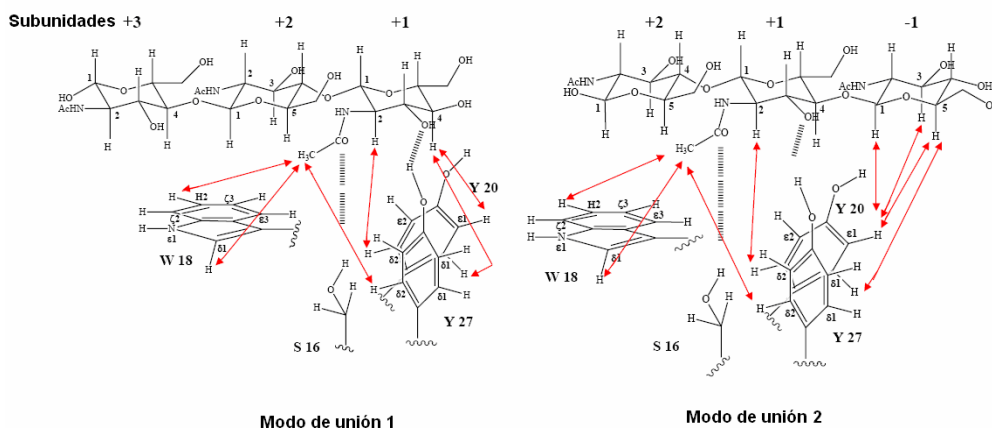


Figura 6.11: Representación de los dos modos de unión en la mutante F18W.

En la tabla siguiente quedan recogidos los NOEs intermoleculares encontrados. Estos NOEs sólo se pueden explicar con la participación de ambos modos de unión.

Protón AcAMP2-F18W	δ (ppm) AcAMP2-F18W	Protón Carbohidrato	δ (ppm) Carbohidrato	Modo de unión A subunidades +3, +2, +1	Modo de unión B subunidades +2, +1, -1
18 H δ 1	7.17	CH ₃	1.96	+1	+1
18 HH2	7.15	CH ₃	1.96	+1	+1
20 H δ 1,2	7.01	H1	4.48	-	-1
20 H δ 1,2	7.01	H2	3.78	+1	+1
20 H δ 1,2	7.01	H3	3.57	-	-1
20 H δ 1,2	7.01	H4	3.43	+1	-
20 H δ 1,2	7.01	H5	3.48	-	-1
20 H ϵ 1,2	6.84	H1	4.48	-	-1
20 H ϵ 1,2	6.84	H3	3.57	-	-1
20 H ϵ 1,2	6.84	H4	3.43	+1	-
20 H ϵ 1,2	6.84	H5	3.48	-	-1
27 H δ 1,2	7.2	CH ₃	1.96	+1	+1

Tabla 6.5: NOEs intermoleculares observados para la mutante F18W. Los NOEs marcados en rojo solo pueden ser explicados mediante el modo de unión 1 mientras que los marcados en azul solo se pueden explicar con el modo de unión 2.

A continuación se muestra una visión del sitio de reconocimiento de azúcares calculada para un modelo del complejo de *Ac-AMP2-F18W*/(GlcNAc)₃ según el modo de unión 1 (**Figura 6.12**).

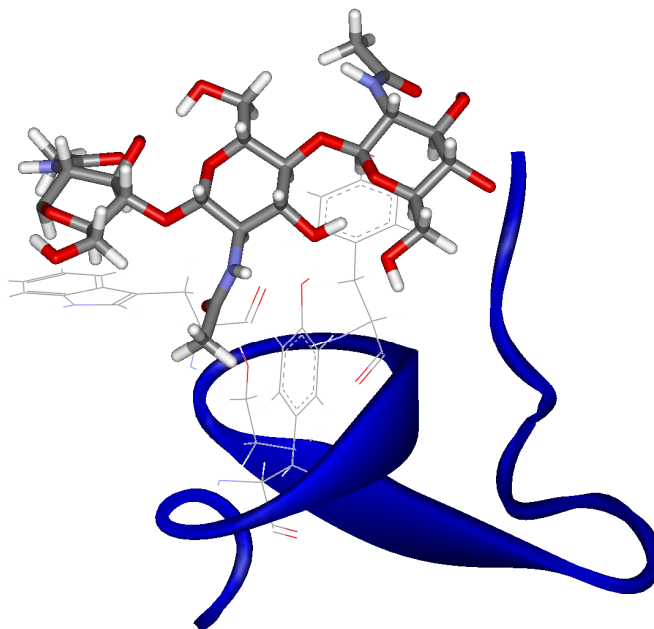


Figura 6.12: Representación esquemática de las interacciones de apilamiento y de enlace de hidrógeno en el sitio de reconocimiento de azúcares del mutante de *Ac-AMP2-F18W*.

3. Determinación de la estructura 3D del mutante Ac-AMP2-F18Nal

La proteína mutante de Ac-AMP2-F18Nal tiene la peculiaridad de poseer un aminoácido no natural en la posición 18. Se trata de la β -2-naftil alanina que, para mayor facilidad, se ha abreviado como Nal.

3.1. Análisis de los espectros

Combinando la información obtenida de los distintos espectros, se consiguió la asignación de las señales de resonancia de protón del péptido (Tabla 6.6 y figura 6.13)

	HN	α H	β H	γ H	Otros
Val1		3.93	2.26	1.09	
Gly2	8.99	4.07, 4.11			
Glu3	8.73	4.44	1.96, 2.07	2.31	
Cyss4	7.84	4.49	2.83, 3.17		
Val5	8.43	4.15	1.93	0.91	
Arg6	9.39	3.92	1.86, 1.94	1.62	H δ : 3.21
Gly7	8.37	3.75, 4.15			
Arg8	7.92	4.77	1.87, 1.93	1.59, 1.68	H δ : 3.21
Cyss9	8.67	5.24	2.32, 2.88		
Pro10		4.47	1.74, 2.41	1.98, 2.03	H δ : 3.62, 4.06
Ser11	8.39	4.17	3.86		
Gly12	8.94	3.7, 4.14			
Met13	7.98	4.65	1.77, 2.06	2.36	
Cyss14	9.28		2.44, 3.35		
Cyss15	8.83	4.86	2.93		
Ser16	9.82	4.95	4.37		
Gln17	9.16	4.06	1.54, 1.74	1.04, 1.46	
Nal18	7.51	4.88	3.08, 3.82		Aromáticos: 7.44, 7.51, 7.54, 7.58, 7.86, 7.90, 7.98
Gly19	7.91	3.62, 4.06			
Tyr20	7.51	5.12	2.90, 3.32		H δ : 7.04, H ϵ : 6.90
Cyss21	8.87	5.56	2.85		
Gly22	8.49	1.92, 3.71			
Lys23	8.11	5.02	1.70, 2.07	1.38, 1.47	H ϵ : 2.93
Gly24	8.42	4.00, 4.60			
Pro25		4.32	1.99, 2.35	2.10	H δ : 3.68, 3.88
Lys26	8.87	4.09	1.66, 1.86	1.45, 1.21	
Tyr27	7.66	4.16	2.56, 2.99		H δ : 7.21, H ϵ : 6.77
Cyss28	8.53	4.32	2.78, 3.26		
Gly29	8.05	4.03			
Arg30	8.35	4.32	1.75, 1.88	1.61, 1.72	H δ : 3.20, H ϵ : 7.14

Tabla 6.6: Desplazamientos químicos (δ) de ^1H -RMN del Ac-AMP2-F18Nal (25° C, pH 5,6).

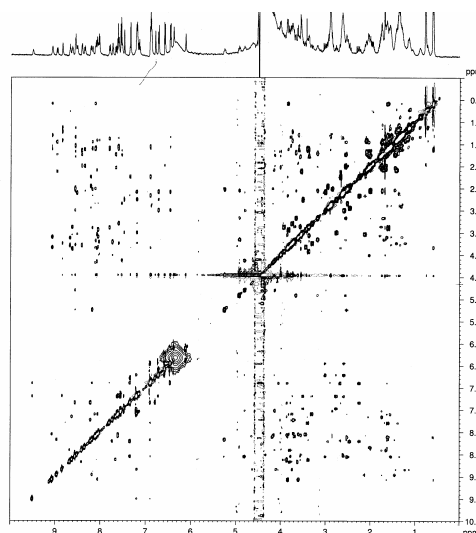


Figura 6.13: Espectro NOESY (500 MHz) del Ac-AMP2-F18Nal en H₂O:D₂O (90:10) y tiempo de mezcla 200 ms.

3.1.1. Protocolo de identificación de estructuras secundarias:

Se asignaron 348 NOEs, además de 67 NOEs no útiles desde el punto de vista estructural. Estos NOEs se pueden clasificar en: 64 intra-residuos, 140 secuenciales, 52 de rango medio ($r < 5$) y 99 de rango largo ($r > 5$). Esta clasificación viene representada en las **Figuras 6.14** y **6.15**, donde se relacionan la distribución de restricciones en función de su rango y en función del tipo de residuo, respectivamente.

En este polipéptido mutante también se pudo determinar la existencia de una lámina β , que se extiende desde la Met13 hasta la Lys23. Además de observar varios $H\alpha$ con desplazamientos químicos mayores que los esperados para un ovillo estadístico, también se observaron una serie de NOEs característicos de lámina β : $H\alpha_{15}$ -HN16, $H\alpha_{21}$ -HN22 (tipo αH_i -NH $_{i+1}$), $H\alpha_{15}$ - $H\alpha_{21}$ (tipo αH_i - αH_j) y HN14- $H\alpha_{23}$, HN15- $H\alpha_{21}$, (tipo NH $_i$ - αH_j).

Aminoácido	Valor Experimental	Valor teórico
Met 13	4.65	4.52
Ser 16	4.97	4.50
Gly 19	3.62/4.06	3.97
Tyr 20	5.12	4.60
Lys23	5.02	4.36

Tabla 6.7: Valores experimentales y esperados para un ovillo estadístico de algunos $H\alpha$ del Ac-AMP2-F18Nal, implicados en la formación de la lámina β .

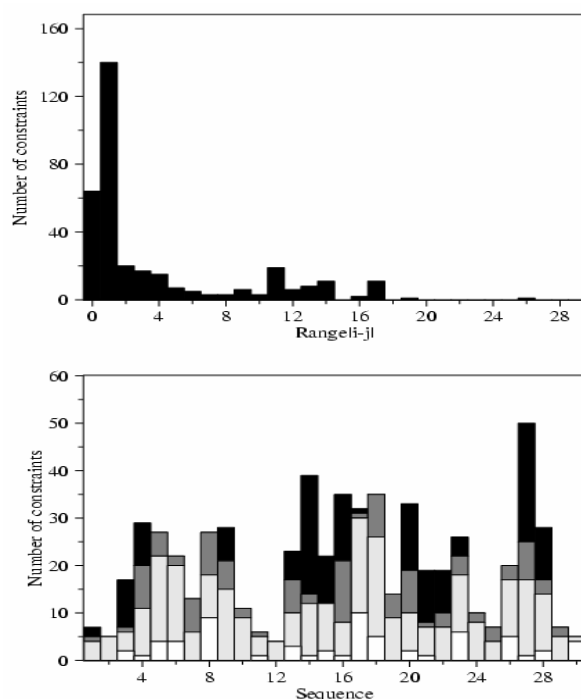


Figura 6.14 y 6.15: Distribución de las restricciones halladas por RMN. En la Figura 6.14 podemos ver el número de restricciones que hay dependiendo del tipo: secuenciales, a media distancia y a larga distancia. En la Figura 6.15 se representa el tipo de restricciones que hay por aminoácido.

3.1.2. Protocolo de identificación de estructuras secundarias:

El procedimiento que se llevó a cabo con esta mutante fue similar al realizado anteriormente. Los NOEs obtenidos se tradujeron a restricciones de distancia y se introdujeron en el programa DYANA. Se realizaron dos cálculos, uno con 400 estructuras iniciales y otro con 800. De cada uno se escogieron las mejores 25 estructuras en función de su función diana. En la **Tabla 6.8** se recogen todos los datos obtenidos y a continuación se representan las figuras de los distintos conjuntos de estructuras (**Figura 6.16**). Los datos indican una buena definición del polipéptido, tanto para su cadena principal ($\text{RMSD} < 0.65 \text{ \AA}$), como para todos los átomos ($\text{RMSD} < 1.3 \text{ \AA}$).

Nº Estructuras	Rango superposición	RMSD (\AA) Cadena principal	RMSD (\AA) Átomos pesados	Función dyana
25	4-28	0.61 ± 0.15	1.17 ± 0.17	0.026 (0.012-0.036)
	13-23(Lámina β)	0.47 ± 0.12	1.03 ± 0.18	

Tabla 6.8: Datos estadísticos y bondad de las estructuras obtenidas mediante los cálculos DYANA para el mutante *Ac-AMP2-F18Nal* ($\text{RMSD} (\text{\AA})$ es la desviación cuadrática media).

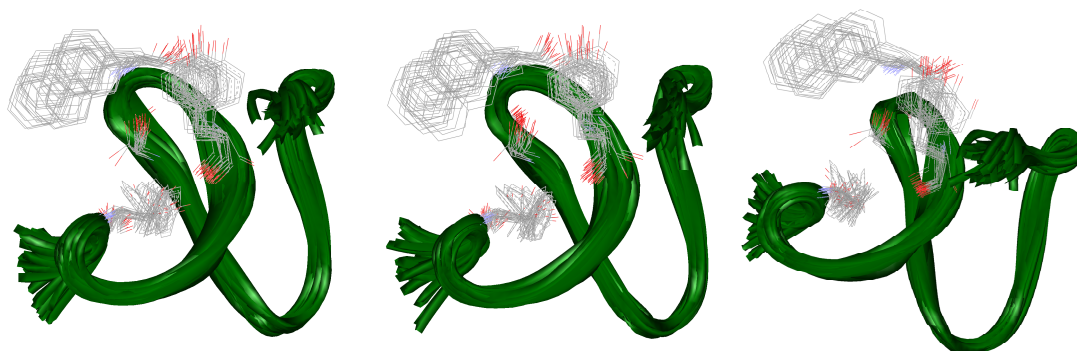


Figura 6.16: Representación esquemática de las estructuras (esqueleto y residuos implicados en el reconocimiento de quitoooligosacáridos) del mutante *Ac-AMP2-F18Nal*, obtenidas mediante DYANA:

- a) Las 25 mejores estructuras del cálculo A
- b) Las 25 mejores estructuras del cálculo B
- c) Superposición de las 50 estructuras seleccionadas.

Tras los cálculos con DYANA, las 50 estructuras obtenidas se refinaron, como en el caso anterior, usando un protocolo de dinámica molecular con restricciones experimentales (RMD) y el campo de fuerzas AMBER. Así, se obtuvo un conjunto de estructuras de las que se escogieron las 25 mejores, en función de su RMSD. En la **Figura 6.17** se representa la superposición de las 25 estructuras seleccionadas mientras que en la **Tabla 6.9** se recogen los datos obtenidos tras la aplicación de este protocolo.

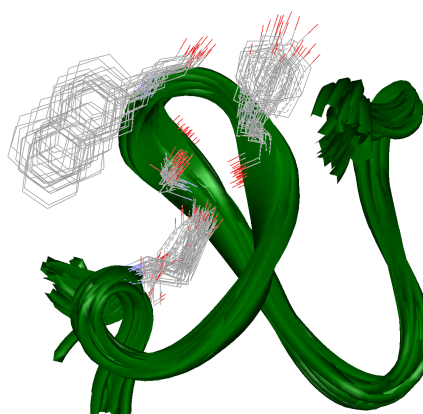


Figura 6.17: Conjunto de estructuras Amber: selección de las 25 mejores estructuras según su RMSD.

Dinámica Molecular (Amber) <i>25 estructuras</i>	Rango superposición	RMSD (Å) Cadena principal	RMSD (Å) Átomos pesados
	4-27	0.81±0.18	1.45±0.27
	13-23(Lámina β)	0.42±0.13	1.87±0.45

Tabla 6.9: Datos estadísticos y bondad de las estructuras obtenidas mediante cálculos AMBER para el mutante *Ac-AMP2-F18Nal* (RMSD (Å) es la desviación cuadrática media).

A continuación, se representa el mapa de Ramachandran de las 25 estructuras seleccionadas de los cálculos con AMBER (**Figura 6.18**). Los puntos representan las glicinas, mientras que cruces representan el resto de aminoácidos.

Puede observarse que prácticamente todos los valores ϕ , ψ se encuentran en zonas permitidas, salvo algunas glicinas (que más móviles desde el punto de vista conformacional) y algunos residuos de algunas las estructuras generadas, que están localizados en los extremos *N*-terminal y *C*-terminales, normalmente correspondientes a las regiones más flexibles en péptidos y proteínas.

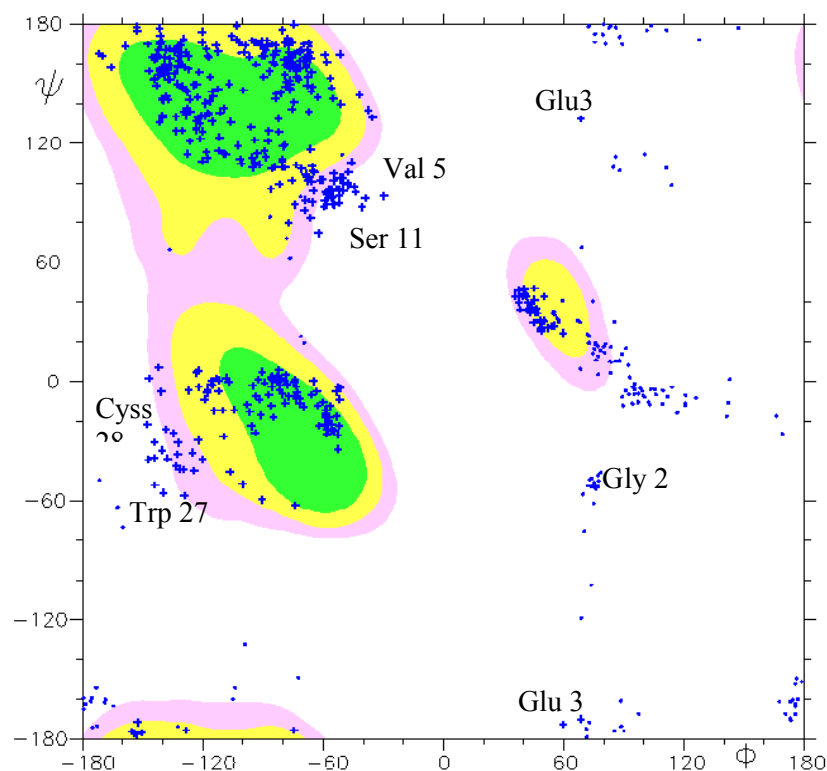


Figura 6.18: Mapa de Ramachandran de las 25 estructuras AMBER del mutante *Ac-AMP2-F18Nal*.

3.2. Análisis de la proteína asociada a (GlcNAc)₃

De igual modo que en el caso de mutante con triptófano, una vez obtenida la estructura 3D del péptido en estado libre, se procedió al estudio de su interacción con *N,N',N''*-triacetilquitotriosa (GlcNAc)₃. La existencia de interacción se dedujo a partir de las variaciones de desplazamientos químicos (δ) observadas en los protones de la proteína, al añadir cantidades crecientes de azúcar (véase más adelante) donde los mayores variaciones se observaron de nuevo en torno a los residuos 16-28

Con objeto de obtener la estructura del complejo Ac-AMP2-F18NaI con quitotriosa, se realizaron espectros NOESY para una relación molar 4:1 entre (GlcNAc)₃ y proteína. Como en el caso anterior, los NOEs detectados pusieron de manifiesto que no existían cambios significativos en el número ni en la intensidad de los picos de cruce de los protones de la proteína en presencia del azúcar, con respecto a los mismos en la proteína libre. Este hecho experimental implica que no existen cambios conformacionales notables en el esqueleto de la proteína al pasar del estado libre al asociado. Además, la observación de NOEs intermoleculares confirmó la misma posición del sitio de reconocimiento (**Figura 6.19**).

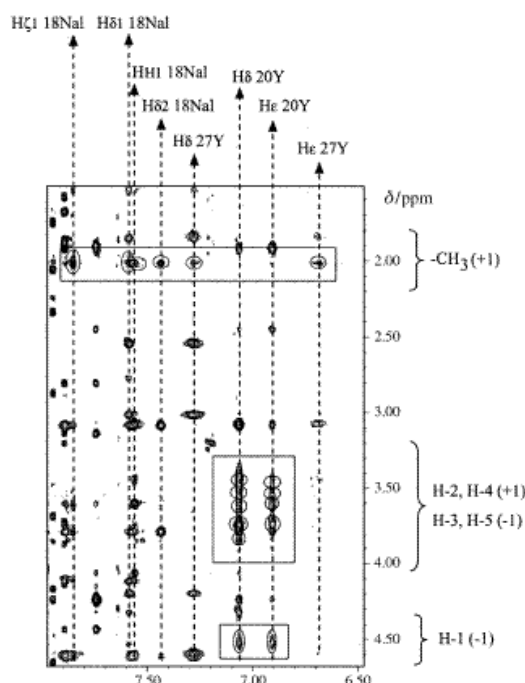


Figura 6.19: Detalle del espectro NOESY con algunos de los NOEs proteína-ligando observados.

En la **Figura 6.20** se reflejan los dos posibles modos de unión entre la proteína y el trisacárido, así como los NOEs posibles en cada uno de ellos. De acuerdo con

los NOEs encontrados (**Figura 6.19** y **Tabla 6.10**) ambos modos de unión tienen lugar en disolución.

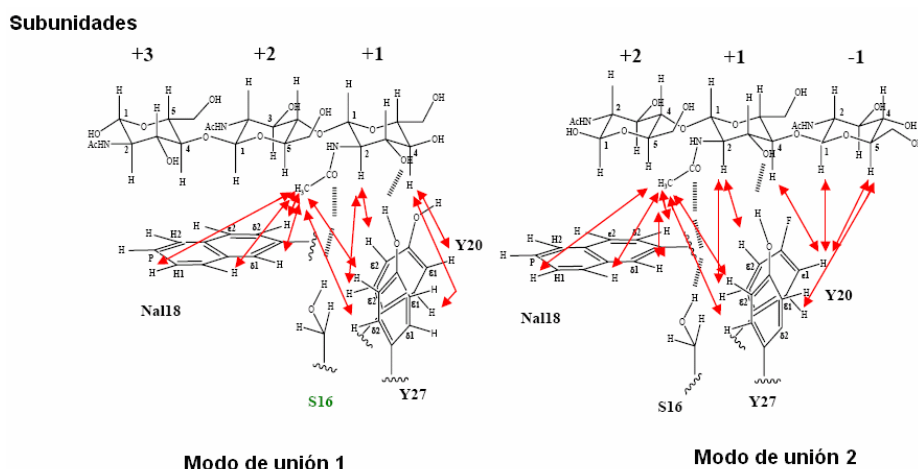


Figura 6.20: Representación de los dos modos de unión en el mutante F18Nal.

Protón AcAMP2-F18Nal	δ (ppm) AcAMP2- F18Nal	Protón Carbohidrato	δ (ppm) Carbohidrato	Modo de unión A subunidades +3, +2, +1	Modo de unión B subunidades +2, +1, -1
18 H δ 1	7.57	CH ₃	2.01	+1	+1
18 H δ 2	7.43	CH ₃	2.01	+1	+1
18 H ϵ 1	7.85	CH ₃	2.01	+1	+1
18 H η 1	7.55	CH ₃	2.01	+1	+1
20 H δ 1,2	7.06	H1	4.51	-	-1
20 H δ 1,2	7.06	H2	3.73/ 3.82	+1	+1
20 H δ 1,2	7.06	H3	3.57	-	-1
20 H δ 1,2	7.06	H4	3.48/ 3.62	+1	+1
20 H δ 1,2	7.06	H5	3.5	-	-1
20 H ϵ 1,2	6.91	H1	4.51	-	-1
20 H ϵ 1,2	6.91	H2	3.73/ 3.82	+1	+1
20 H ϵ 1,2	6.91	H3	3.57	-	-1
20 H ϵ 1,2	6.91	H4	3.48/ 3.62	+1	+1
20 H ϵ 1,2	6.91	H5	3.5	-	-1
27 H δ 1,2	7.28	CH ₃	2.01	+1	+1
27 H ϵ 1,2	6.69	CH ₃	2.01	+1	+1

Tabla 6.10: NOEs intermoleculares observados para la mutante F18Nal. Los NOEs marcados en rojo sólo pueden ser explicados mediante el modo de unión 1, mientras que los marcados en azul solo se pueden explicar con el modo de unión 2.

A continuación se muestra una visión del sitio de reconocimiento de azúcares calculada para un modelo del complejo de Ac-AMP2-F18Nal/(GlcNAc)₃ según el modo de unión 1 (**Figura 6.21**).

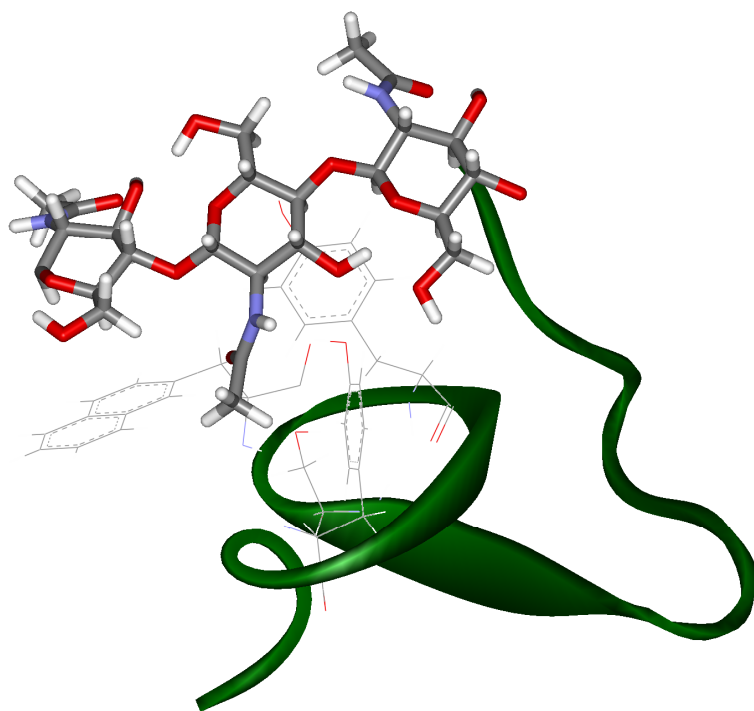


Figura 6.21: Representación esquemática de las interacciones de apilamiento y de enlace de hidrógeno en el sitio de reconocimiento de azúcares del mutante de Ac-AMP2.

4. Comparativa entre las mutantes y la proteína nativa Ac-AMP-2:

La estructura de la proteína nativa³⁰ es similar a las encontradas en este trabajo. El RMSD del esqueleto entre las estructuras de RMN de Ac-AMP2, Ac-AMP2-F18W y Ac-AMP2-F18Nal es de solamente 0.95 Å. De hecho, existe también la lámina β anti-paralela, que se extiende desde la Met 13 a la Lys 23. Además, tal y como puede observarse en la **Figura 6.22**, el Ac-AMP2 presenta sus giros en las mismas posiciones que los péptidos mutantes. Es decir, desde un punto de vista estructural, la mutación puntual en la posición 18 no afecta a la conformación global del polipéptido. Ni la mutación F18W ni la F18Nal suponen un cambio conformacional importante respecto a la estructura nativa del Ac-AMP2, tanto a escala de la cadena principal, como de la orientación de las cadenas laterales. Por tanto, ambos mutantes son modelos adecuados para estudiar la importancia de la naturaleza de los anillos aromáticos de lectinas para el reconocimiento de hidratos de carbono.

RMSD cadena principal (3-27) = 0.95

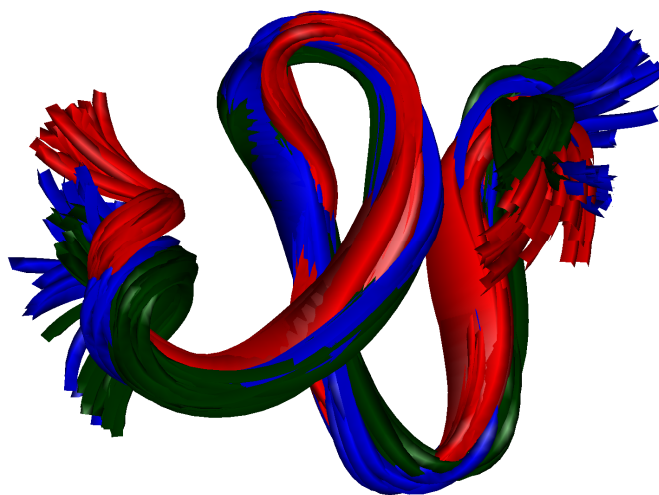
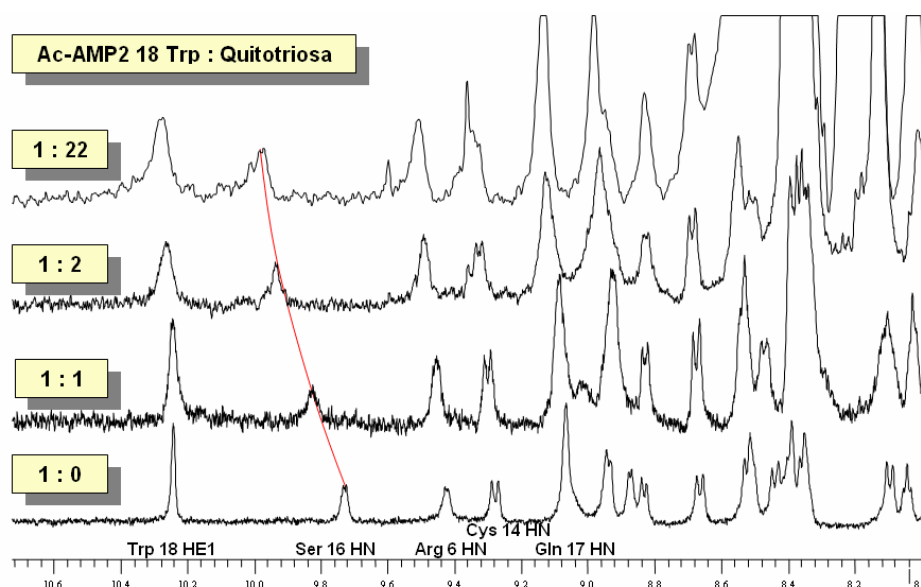


Figura 6.22: Superposición de la estructura descrita para el Ac-AMP2 mediante RMN (rojo) con las obtenidas en este trabajo para sus mutantes F18W (azul) y F18Nal (verde).

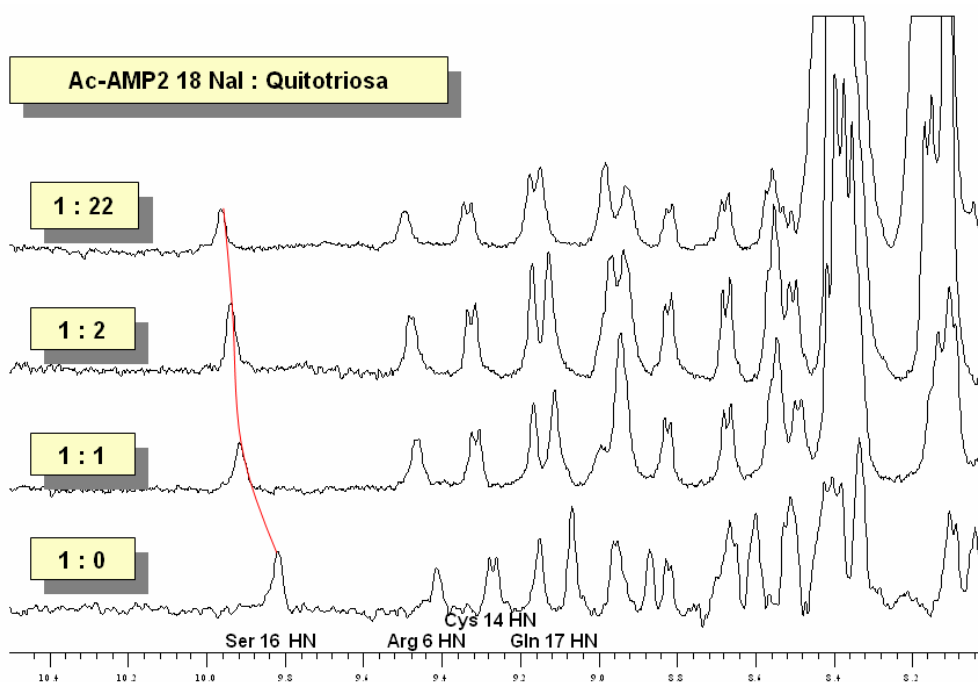
5. Estudios de la interacción entre Ac-AMP2 y sus mutantes con(GlcNAc)₃.

Una vez determinadas las estructuras 3D de ambos mutantes, se procedió al estudio de su interacción con el ligando *N,N',N''*-triacetilquitotriosa, que contiene la unidad GlcNAc. Además, se realizó el mismo estudio con la proteína nativa, Ac-AMP2.

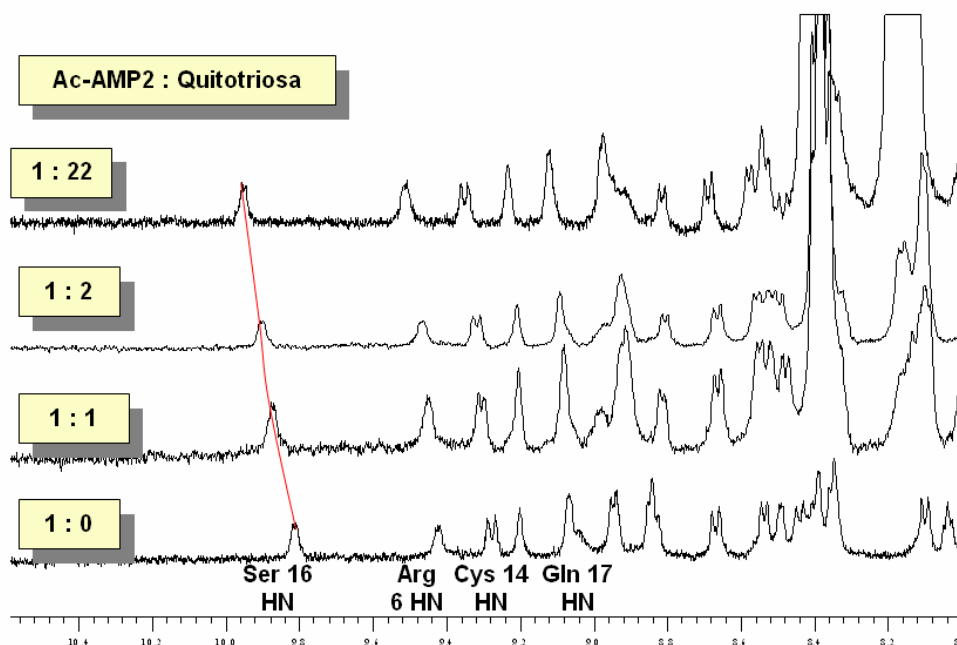
Este estudio se realizó mediante valoraciones, utilizando la metodología de RMN puesta a punto en nuestro laboratorio^{8,42}. En las tres valoraciones llevadas a cabo se mantuvo constante la concentración de proteína y se varió la concentración de azúcar. Con el fin de obtener los parámetros termodinámicos de van't Hoff, estos experimentos se llevaron a cabo a 4 temperaturas distintas. En los tres casos, en la proteína nativa, su mutante F18Nal y su mutante F18W se siguió la señal de protón correspondiente al NH de la Ser 16. Estas señales resultaron ser muy adecuadas, por encontrarse en una zona muy aislada del espectro, además de ser de aquellas que más variación de δ experimentan al aumentar la concentración de azúcar. En la **Figura 6.22**, **Figura 6.23** y **Figura 6.24** se representa los espectros realizados a 25°C en distintos puntos de la valoración para cada una de las proteínas (Ac-AMP2-F18W, Ac-AMP2-F18Nal y Ac-AMP2, respectivamente).



6.22: Diferentes espectros de ¹H-RMN (500 MHz, 25 °C, pH 5.6) tomados en cuatro puntos de la valoración con (GlcNAc)₃, de una disolución de Ac-AMP2-F18W. Puede observarse el desplazamiento sistemático de la señal del NH de la Ser 16, en función de la concentración de trisacárido en el tubo de RMN.



6.23: Diferentes espectros de ¹H-RMN (500 MHz, 25 °C, pH 5.6) tomados en cuatro puntos de la valoración con (GlcNAc)₃, de una disolución de Ac-AMP2-F18Nal. Puede observarse el desplazamiento sistemático de la señal del NH de la Ser 16, en función de la concentración de trisacárido en el tubo de RMN.



6.24: Diferentes espectros de ¹H-RMN (500 MHz, 25 °C, pH 5.6) tomados en cuatro puntos de la valoración con (GlcNAc)₃, de una disolución de Ac-AMP2. Puede observarse el desplazamiento sistemático de la señal del NH de la Ser 16, en función de la concentración de trisacárido en el tubo de RMN.

Si se representan los desplazamientos químicos observados frente a la concentración de quitotriosa en cada punto, se obtienen las gráficas de las valoraciones de las tres proteínas (**Figura 6.25**). El ajuste a la ecuación correspondiente permite obtener los datos de constantes de asociación (**Tabla 6.11**).

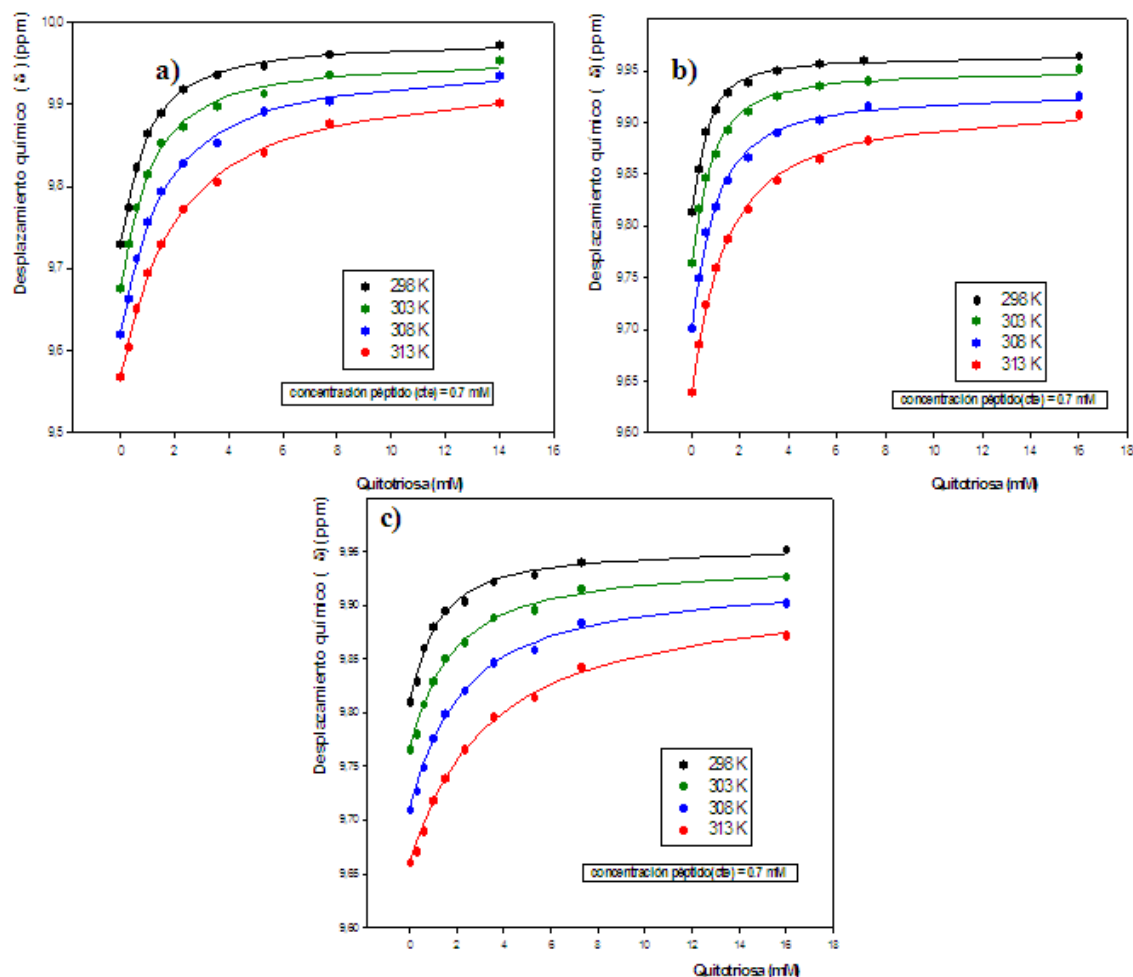


Figura 6.25: Variación del δ de ^1H -RMN del NH de la Ser16 con la concentración de $(\text{GlcNAc})_3$ en función de la temperatura en el Ac-AMP2-F18W (a), en el Ac-AMP2-F18Nal (b) y en la proteína nativa (c).

	Constante de afinidad (M^{-1}) (K_a)				Parámetros termodinámicos	
	298 K	303 K	308 K	313 K	ΔH (KJ/mol)	ΔS (J/mol)
Ac-Amp2 F18W	1755	1473	916	641	-54.1	-119.0
Ac-Amp2 F18Nal	3527	2314	1638	999	-63.9	-146.6
Ac-Amp2	1206	721	537	340	-63.4	-154.1

Tabla 6.11: Valores de las constantes de asociación y de los parámetros termodinámicos de los procesos de asociación intermolecular entre $(\text{GlcNAc})_3$ y los receptores Ac-AMP2 y sus mutantes F18W y F18Nal.

Una vez obtenidos los valores de K_a para cada temperatura, se determinaron los parámetros termodinámicos de asociación ΔH° y ΔS° mediante una representación de van't Hoff. En la **Figura 6.26** se representan los puntos y los ajustes correspondientes.

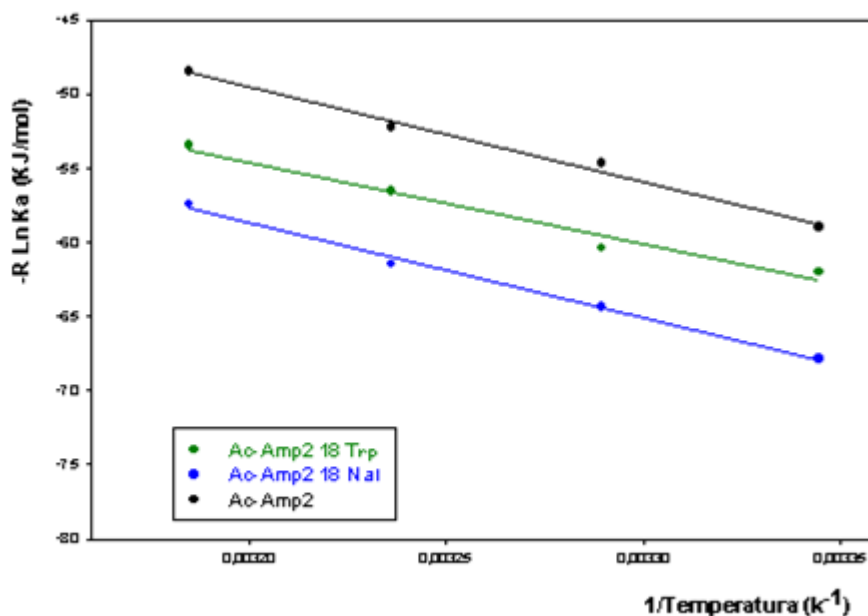


Figura 6.26: Representación de Van't Hoff de la que se obtienen los parámetros termodinámicos de la valoración.

Los resultados indican que, en los tres casos, el proceso de interacción del AcAMP2 (o sus mutantes) con el (GlcNAc)₃ está entálpicamente favorecido y entrópicamente desfavorecido, como ya ha sido reportado en otros dominios de heveína^{8,27,43}. Es importante señalar que la obtención de los parámetros termodinámicos a partir de este tipo de análisis asume que la capacidad calorífica no depende de la temperatura, hecho que no es siempre cierto en interacciones entre biomoléculas. Por tanto, las magnitudes de los valores de ΔH° y ΔS° deben ser consideradas solamente cualitativamente y deben ser comparados con las obtenidos mediante otras técnicas más directas, como microcalorimetría. Estudios previos con dominios de heveína ha demostrado que los valores de RMN difieren en menos de un 10% a los obtenidos por otras técnicas, como microcalorimetría.

Aunque el origen del fenómeno de compensación entalpía-entropía sigue en cuestión, se ha publicado para magnitudes y signos de entropía y entalpía similares que los enlaces de hidrógeno y fuerzas de van der Waals deben ser los factores predominantes en la estabilización del complejo. Parte de la entropía negativa observada podría ser debida a la rigidificación del carbohidrato y/o de las cadenas laterales de la proteína. Como se dedujo en estudios anteriores^{5,8,10-11,13,26,44-45}, la máxima pérdida de entropía conformacional por congelación del ligando en la interacción con los dominios de heveína alcanza los 17 kJ/mol a 25°C, mientras que la observada en estos mutantes oscila entre 30 y 46 KJ/mol. Este hecho implica que además de la rigidificación del trisacárido se produce una pérdida entrópica adicional que parece ser debida a la reducción de flexibilidad en las cadenas laterales de estos dominios de heveína, propiciada probablemente a su menor tamaño y a la necesidad de albergar el trisacárido, obteniendo a su vez una entalpía de enlace favorable.

Por otro lado, los valores de K_a para la proteína nativa y sus mutantes indican que, cuanto mayor es la superficie del aromático en la posición 18, mayor es la K_a . Por tanto, la interacción carbohidrato:proteína mejora con un aumento de la densidad electrónica del anillo aromático involucrado en la interacción CH- π . Estudios anteriores de nuestro grupo⁴⁶ han puesto de manifiesto que la introducción de átomos de flúor en los anillos aromáticos clave de la cadena de AcAMP2 producen una disminución notable de la afinidad. Es decir, la naturaleza química del anillo aromático, y su grado de activación, influye decisivamente en la magnitud del proceso de reconocimiento molecular.

Las interacciones descritas, tanto desde un punto de vista estructural como energético, para estos dominios se han observado también en las proteínas del tipo heveína estudiadas anteriormente por nuestro grupo de trabajo. Tanto interacciones de apilamiento, van der Waals y polares contribuyen al proceso de asociación, estabilizando los complejos a través de la formación de enlaces de hidrógeno y de interacciones de apilamiento. Por tanto, la visión estructural obtenida en disolución a través de RMN y modelización está de acuerdo con el carácter entálpicamente dirigido de la asociación, deducida de la medida de los parámetros termodinámicos.

6. Conclusiones:

El estudio de las estructuras tridimensionales de los mutantes F18W y F18Nal del Ac-AMP2, mediante RMN, indica que ambos dominios presentan plegamientos bien definidos, que pertenecen al tipo heveína. Ni la mutación F18W ni la F18Nal producen un cambio conformacional importante respecto a la estructura tridimensional nativa del Ac-AMP2, tanto a escala de la cadena principal, como de la orientación de las cadenas laterales. Por tanto, ambos mutantes son modelos adecuados para estudiar la importancia de la naturaleza de los anillos aromáticos de lectinas para el reconocimiento de carbohidratos.

Por otro lado, las estructuras de las dos proteínas mutantes libres y asociadas a (GlcNAc)₃ son similares, lo que indica que el sitio de reconocimiento de quitooligosacáridos de estas proteínas está preorganizado.

Finalmente, los datos termodinámicos de asociación proteína-carbohidrato indican que el proceso de reconocimiento molecular está entálpicamente dirigido y que tanto los enlaces de hidrógeno como las interacciones de Van der Waals contribuyen a la estabilidad del complejo. Además, el papel clave de las interacciones CH- π se ha puesto de manifiesto en este estudio al observar un aumento de la K_a a medida que se aumenta la superficie del anillo aromático.

7. Parte Experimental:

7.1. Preparación de las muestras y experimentos de RMN:

Todas las muestras se prepararon en tampón fosfato 20mM, 100mM NaCl, a pH 5.6. En el caso del Ac-AMP2-F18W en estado libre, se usó una muestra de concentración de 2.76 mM mientras que la concentración del Ac-AMP2-F18Nal fue de 2.92 mM.

En estas condiciones, y a 25°C, se realizaron experimentos TOCSY⁴⁷ y NOESY⁴⁸ en espectrómetros Bruker DRX y AMX de 500 MHz usando la secuencia Watergate para la supresión del agua. Para el Ac-AMP2-F18W libre se realizó un TOCSY, con un tiempo de mezcla de 60ms y dos NOESYs con tiempos de mezcla de 150 y de 300 ms mientras que para el Ac-AMP2-F18Nal libre se realizó un experimento NOESY con 200ms de tiempo de mezcla.

En el caso de los complejo con *N,N',N''*-triacetilquitotriosa, se usó una muestra con concentración del péptido de 1.98 mM y 16mM de azúcar (ca. 8:1) en ambos casos. En el caso del Ac-AMP2-F18W, se realizaron experimentos en espectrómetros Bruker AMX de 500 MHz y Avance de 800 MHz. El TOCSY se obtuvo con tiempo de mezcla de 60 ms, mientras que los NOESYs se registró con tiempos de mezcla de 200 y 300 ms. Para el Ac-AMP2-F18Nal se realizó un TOCSY con tiempo de mezcla de 60 ms y un NOESY con tiempo de mezcla de 200 ms en el en espectrómetros Bruker AMX de 500 MHz. También se realizó un experimento NOESY a 800 MHz con un tiempo de mezcla de 200 ms.

7.2. Construcción de la librería del aminoácido no natural

Tanto el programa XEASY, como los programas DYANA y AMBER, poseen librerías con las estructuras y cargas de cada aminoácido. Sin embargo, en el caso del mutante AcAMP2-F18Nal, con un aminoácido no natural, fue necesario crear una librería adecuada para este residuo. El proceso de creación de una librería tiene varios pasos. El primero consiste en crear la estructura del aminoácido y optimizarlo. Esto se realizó usando el programa Chem3D Pro con el campo de fuerzas MM2. Una vez minimizada la estructura, el segundo paso fue la determinación de las cargas, utilizando Gaussian 98⁴⁹, mediante el método Hartree Fock 6-31G. Se utilizó

este método porque es el que se había empleado para la obtención de las librerías del resto de aminoácidos.

Para obtener la librería de DYANA o de XEASY, fue necesario introducir el fichero PDB del aminoácido en el programa MOLMOL, que ya es capaz de crear las librerías adecuadas. Los programas DYANA³⁸, XEASY³⁷ y MOLMOL⁵⁰ han sido creados por el grupo del Prof. Wüthrich, en Zürich, y están disponibles comercialmente.

En el caso de la librería para el programa AMBER⁴⁰ fue necesario, además, introducir las cargas. Para ello se usó el módulo XLEAP¹⁵ de AMBER que es capaz de editar y modificar el fichero PDB correspondiente.

7.3. Generación de estructuras 3D: DYANA y RMD (AMBER):

El cálculo de las estructuras 3D se inició mediante un protocolo de dinámica en el espacio de los ángulos de torsión con el programa DYANA. Las restricciones de distancia superiores usadas en DYANA fueron obtenidas mediante las intensidades de los picos de cruce NOESY, analizados a través del programa XEASY. Estas intensidades se clasificaron como fuertes, medias y débiles, y se transformaron en límites superiores de 2.5, 3.5 y 5.5 Å, respectivamente. Las restricciones de límites inferiores se colocaron de acuerdo a la suma de radios van der Waals entre dos átomos. Las restricciones para los protones diastereotópicos no asignados estereoespecíficamente se definieron de acuerdo a una posición intermedia (pseudoátomo), adicionando a la distancia 2.20 Å.

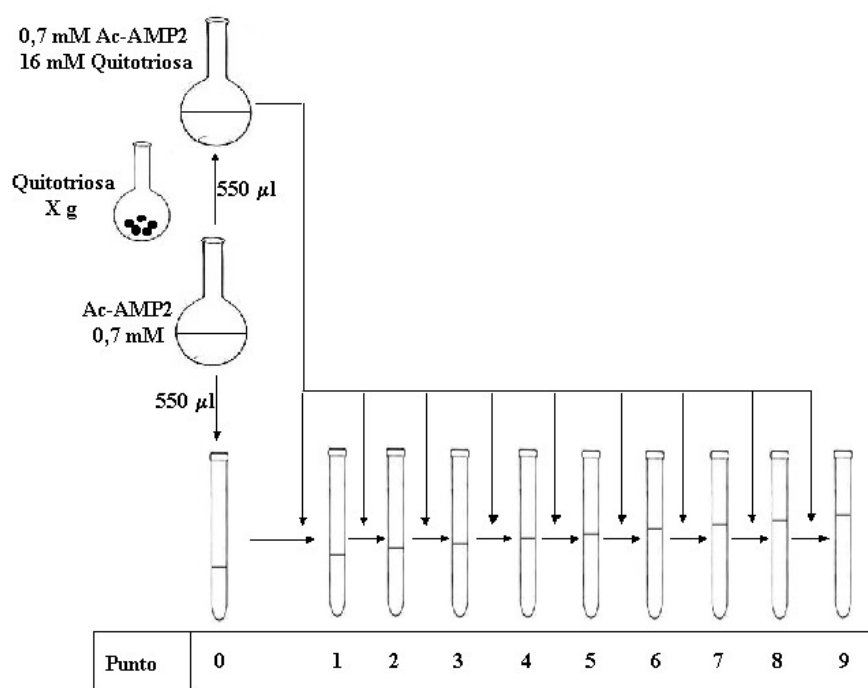
Los enlaces disulfuro se incluyeron como restricciones de distancia entre S-S ($2.0 < r < 2.1$ Å) y para C β -S ($3.0 < r < 3.1$ Å).

Se realizaron dos cálculos, uno con 400 estructuras y un segundo con 800, de las cuales se seleccionaron las mejores 25 estructuras de cada uno de ellos basándose para ello en su función diana. Estas 50 estructuras se sometieron a un protocolo de dinámica molecular restringida (RMD), con el campo de fuerzas de AMBER. La etapa inicial fue una minimización energética restringida con 2000 iteraciones de gradiente conjugado. Posteriormente, en el proceso de enfriamiento simulado (SA), las estructuras se equilibraron a una temperatura de 600 K durante 2 ps. Durante este paso, se experimentan varios cambios conformacionales y, en el siguiente paso de 2 ps, se inicia realmente la simulación mediante dinámica molecular restringida

(RMD). En un paso subsiguiente, la temperatura se disminuye paulatinamente cada 2 ps hasta llegar a 100 K. Cuando esta última temperatura se alcanza, se llevó a cabo otra RMD de 4 ps. Finalmente, las estructuras se sometieron a una minimización energética restringida (REM), mediante 2000 iteraciones de gradiente conjugado.

7.4. Interacción proteína-carbohidratos: Valoraciones con Quitotriosa

Se preparó 1.1 ml de disolución de proteína (0.7 mM) usando el mismo tampón empleado en la determinación estructural (tampón fosfato 20 mM, 100mM NaCl y pH 5.6). Con el fin de mantener constante la concentración de proteína, se llevó a cabo el procedimiento detallado en el **Esquema 6.3**.



Esquema 6.3: Procedimiento experimental utilizado en este trabajo para realizar los experimentos de valoración, usando disoluciones cada vez más concentradas de $(\text{GlcNAc})_3$ y manteniendo constante la concentración de Ac-AMP2.

Para la determinación de los parámetros termodinámicos de van't Hoff, estos experimentos se llevaron a cabo a 4 temperaturas distintas, 298, 303, 308 y 313 K. En los tres casos, en la proteína nativa, su mutante F18Nal,y su mutante F18W se siguió la señal de ^1H -RMN correspondiente al NH de la Ser 16 (**Tabla 6.12**, **Tabla 6.13** y **Tabla 6.14**)

Punto Valoración	Proteína (nM)	Ligando (mM)	δ Obs 25°C	δ Obs 30°C	δ Obs 35°C	δ Obs 40°C
0	0.7	0	9.730	9.675	9.620	9.568
1		0.29	9.774	9.729	9.663	9.604
2		0.58	9.823	9.777	9.712	9.651
3		1.00	9.865	9.815	9.757	9.694
4		1.51	9.890	9.853	9.794	9.730
5		2.33	9.918	9.873	9.828	9.772
6		3.55	9.936	9.898	9.853	9.805
7		5.29	9.947	9.913	9.891	9.842
8		7.29	9.961	9.936	9.904	9.877
9		16	9.972	9.953	9.935	9.901

Tabla 6.12: Datos de desplazamiento químico de la señal del NH de la Ser 16, obtenidos a partir de las valoraciones de la proteína mutante Ac-AMP2-F18W con (GlcNAc)₃. Se indican los valores obtenidos para cada temperatura.

Punto Valoración	Proteína (nM)	Ligando (mM)	δ Obs 25°C	δ Obs 30°C	δ Obs 35°C	δ Obs 40°C
0	0.7	0	9,814	9,764	9,701	9,639
1		0.29	9,854	9,817	9,749	9,685
2		0.58	9,891	9,846	9,794	9,724
3		1.00	9,912	9,870	9,818	9,760
4		1.51	9,929	9,893	9,844	9,787
5		2.33	9,938	9,910	9,866	9,816
6		3.55	9,950	9,925	9,890	9,844
7		5.29	9,956	9,935	9,902	9,864
8		7.29	9,960	9,940	9,915	9,882
9		16	9,964	9,951	9,925	9,908

Tabla 6.13: Datos de desplazamiento químico de la señal del NH de la Ser 16, obtenidos a partir de las valoraciones de la proteína mutante Ac-AMP2-F18Nal con (GlcNAc)₃. Se indican los valores obtenidos para cada temperatura.

Punto Valoración	Proteína (nM)	Ligando (mM)	$\delta\text{Obs } 25^\circ\text{C}$	$\delta\text{Obs } 30^\circ\text{C}$	$\delta\text{Obs } 35^\circ\text{C}$	$\delta\text{Obs } 40^\circ\text{C}$
0	0.7	0	9,810	9,766	9,710	9,660
1		0.29	9,829	9,780	9,727	9,671
2		0.58	9,860	9,808	9,749	9,690
3		1.00	9,880	9,829	9,776	9,718
4		1.51	9,895	9,851	9,799	9,739
5		2.33	9,904	9,866	9,821	9,766
6		3.55	9,922	9,889	9,847	9,796
7		5.29	9,928	9,896	9,859	9,814
8		7.29	9,940	9,916	9,884	9,843
9		16	9,952	9,927	9,902	9,872

Tabla 6.13: Datos de desplazamiento químico de la señal del NH de la Ser 16, obtenidos a partir de las valoraciones de la proteína mutante Ac-AMP2 con $(\text{GlcNAc})_3$. Se indican los valores obtenidos para cada temperatura.

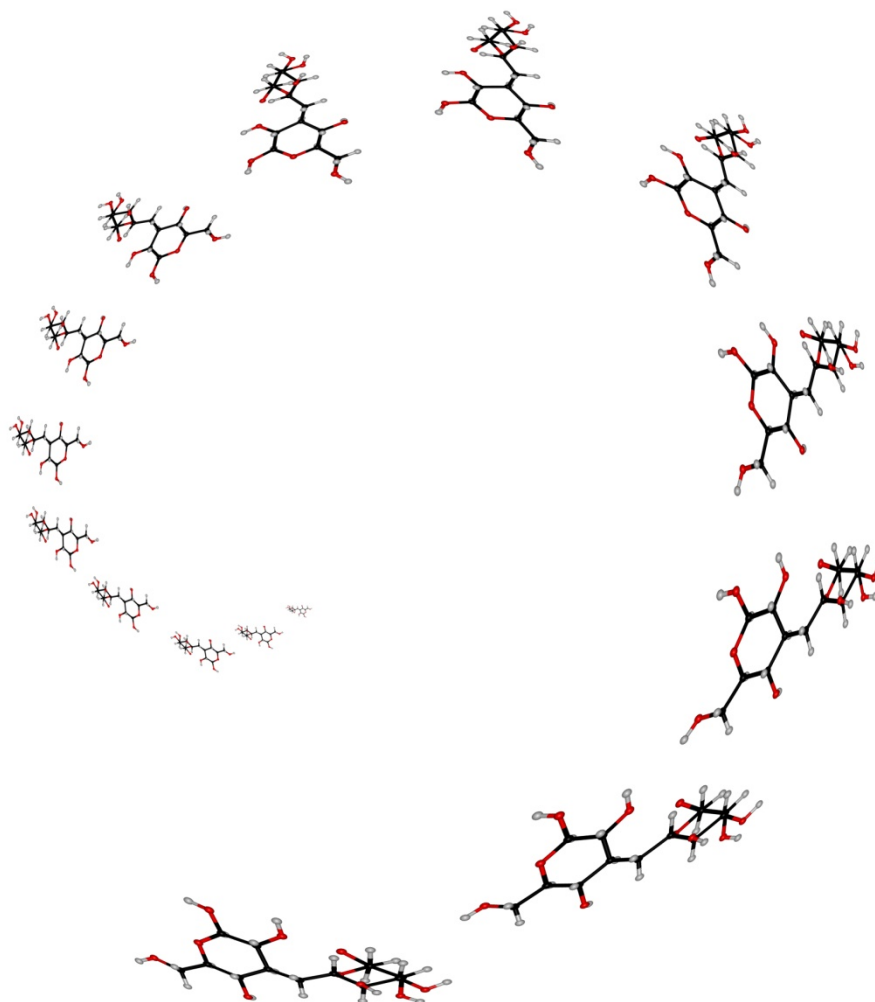
8. Bibliografía:

- (1) Rudiger, H.; Siebert, H. C.; Solis, D.; Jimenez-Barbero, J.; Romero, A.; von der Lieth, C. W.; Diaz-Marino, T.; Gabius, H. J. *Curr Med Chem* **2000**, *7*, 389.
- (2) Van Damme EJM, P. W., Pusztai A, Bardocz S. *Handbook of Plant Lectins: Properties and Biomedical Applications Wiley-VCH*; Wiley-VCH., 1998.
- (3) Raikhel N.V., L. H. I., Broekaert, W.F. ; Mol. Biol. 1993, 591-615. *Annu Rev Plant Physiol Plant Mol Biol* **1993**, *44*, 591.
- (4) Beintema, J. J. *FEBS Lett* **1994**, *350*, 159.
- (5) Rodríguez-Romero, A.; Ravichandran, K. G.; Soriano-García, M. *FEBS Lett* **1991**, *291*, 307.
- (6) Wagner, S.; Breitneder, H. *Biochem Soc Trans* **2002**, *30*, 935.
- (7) Wright C.S. *J Mol Biol* **1984**, *178*, 91.
- (8) Juan Luis Asensio, F. J. C., Marta Bruix, Adela Rodriguez-Romero, Jesús Jimenez-Barbero *Eur J Biochem* **1995**, *230*, 621.
- (9) Poveda, A.; Asensio, J. L.; Espinosa, J. F.; Martin-Pastor, M.; Cañada, J.; Jiménez-Barbero, J. *J Mol Graph Model* **1997**, *15*, 9.
- (10) Asensio, J. L.; Canada, F. J.; Bruix, M.; Gonzalez, C.; Khiar, N.; Rodriguez-Romero, A.; Jimenez-Barbero, J. *Glycobiology* **1998**, *8*, 569.
- (11) Asensio, J. L.; Cañada, F. J.; Siebert, H.-C.; Laynez, J.; Poveda, A.; Nieto, P. M.; Soedjanaamadja, U. M.; Gabius, H.-J.; Jiménez-Barbero, J. *Chem Biol* **2000**, *7*, 529.
- (12) Juan Felix Espinosa, J. L. A., Jose Luis García, José Laynez, Marta Bruix, Christine Wright Hans-Christian, Siebert Hans-Joachim Gabius, Francisco Javier Cañada, Jesus Jiménez-Barbero *Eur J Biochem* **2000**, *267*, 3965.
- (13) Juan Luis Asensio, H.-C. S., Claus-Wilhelm von der Lieth, Jose Laynez, Marta Bruix, U.M. Soedjanaamadja, Jaap J. Beintema, Francisco Javier Canada, Hans-Joachim Gabius, Jesus Jimenez-Barbero *Proteins: Struct, Funct, Genet* **2000**, *40*, 218.
- (14) Poveda, A.; Jimenez-Barbero, J. *Chem Soc Rev* **1998**, *27*, 133.
- (15) Kogelberg, H.; Solís, D.; Jiménez-Barbero, J. *Curr Opin Struct Biol* **2003**, *13*, 646.
- (16) Vyas, N. K. *Curr Opin Struct Biol* **1991**, *1*, 732.
- (17) Nishio M, U. Y., Hirota M, Takeuchi Y, *Tetrahedron* **1995**, *51*, 8665.
- (18) Toone, E. J. *Curr Opin Struct Biol* **1994**, *4*, 719.
- (19) Fernández-Alonso, M. d. C.; Cañada, F. J.; Jiménez-Barbero, J.; Cuevas, G. *J Am Chem Soc* **2005**, *127*, 7379.
- (20) M. Dolores Díaz; María del Carmen Fernández-Alonso; Gabriel Cuevas; F. Javier Cañada; Jiménez-Barbero, J. *Pure Appl Chem* **2008**, *80*, 1827.
- (21) Muraki, M. *Protein Pept Lett* **2002**, *9*, 195.
- (22) Vandenbussche, S.; Díaz, D.; Fernández-Alonso, M. C.; Pan, W.; Vincent, S. P.; Cuevas, G.; Cañada, F. J.; Jiménez-Barbero, J.; Bartik, K. *Chem. Eur. J.* **2008**, *14*, 7570.
- (23) Fernandez-Alonso Mdel, C.; Asensio, J. L.; Canada, F. J.; Jimenez-Barbero, J.; Cuevas, G. *ChemPhysChem* **2003**, *4*, 754.
- (24) Saul, F. A.; Rovira, P.; Boulot, G.; Damme, E. J.; Peumans, W. J.; Truffa-Bachi, P.; Bentley, G. A. *Structure* **2000**, *8*, 593.
- (25) Harata, K.; Muraki, M. *J Mol Biol* **2000**, *297*, 673.

- (26) Andersen, N. H.; Cao, B.; Rodriguez-Romero, A.; Arreguin, B. *Biochemistry (Mosc)* **1993**, *32*, 1407.
- (27) Aboitiz, N.; Vila-Perello, M.; Groves, P.; Asensio, J. L.; Andreu, D.; Canada, F. J.; Jimenez-Barbero, J. *ChemBioChem* **2004**, *5*, 1245.
- (28) Hayashida, M.; Fujii, T.; Hamasu, M.; Ishiguro, M.; Hata, Y. *J Mol Biol* **2003**, *334*, 551.
- (29) Huang, R. H.; Xiang, Y.; Tu, G. Z.; Zhang, Y.; Wang, D. C. *Biochemistry (Mosc)* **2004**, *43*, 6005.
- (30) Martins, J. C.; Maes, D.; Loris, R.; Pepermans, H. A.; Wyns, L.; Willem, R.; Verheyden, P. *J Mol Biol* **1996**, *258*, 322.
- (31) Lemieux, R. U. *Acc Chem Res* **1996**, *29*, 373.
- (32) Archer, B. L. *Biochem J* **1960**, *75*, 236.
- (33) Karisola, P.; Mikkola, J.; Kalkkinen, N.; Airenne, K. J.; Laitinen, O. H.; Repo, S.; Pentikainen, O. T.; Reunala, T.; Turjanmaa, K.; Johnson, M. S.; Palosuo, T.; Kulomaa, M. S.; Alenius, H. *J Immunol* **2004**, *172*, 2621.
- (34) Muraki, M.; Morii, H.; Harata, K. *Protein Eng* **2000**, *13*, 385.
- (35) Muraki, M.; Morii, H.; Harata, K. *Protein Pept Lett* **1998**, *5*, 193.
- (36) de Alba, E.; Santoro, J.; Rico, M.; Jimenez, M. A. *Protein Sci* **1999**, *8*, 854.
- (37) Bartels, C.; Xia, T.-h.; Billeter, M.; Güntert, P.; Wüthrich, K. *J Biomol NMR* **1995**, *6*, 1.
- (38) Güntert, P.; Braun, W.; Wüthrich, K. *J Mol Biol* **1991**, *217*, 517.
- (39) Scheek, R. M.; van Gunsteren, W. F.; Kaptein, R.; Norman, J. O. a. T. L. J. In *Methods Enzymol*; Academic Press: 1989; Vol. Volume 177, p 204.
- (40) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. *J Am Chem Soc* **1984**, *106*, 765.
- (41) Jimenez-Barbero, J.; Asensio, J. L.; Canada, F. J.; Poveda, A. *Curr Opin Struct Biol* **1999**, *9*, 549.
- (42) Jiménez-Barbero, J.; Javier Cañada, F.; Asensio, J. L.; Aboitiz, N.; Vidal, P.; Canales, A.; Groves, P.; Gabius, H.-J.; Siebert, H.-C. In *Adv Carbohydr Chem Biochem*; Derek, H., Ed.; Academic Press: 2006; Vol. Volume 60, p 303.
- (43) Chávez, M. I.; Andreu, C.; Vidal, P.; Aboitiz, N.; Freire, F.; Groves, P.; Asensio, J. L.; Asensio, G.; Muraki, M.; Cañada, F. J.; Jiménez-Barbero, J. *Chem. Eur. J.* **2005**, *11*, 7060.
- (44) Siebert, H.-C.; André, S.; Asensio, J. L.; Cañada, F. J.; Dong, X.; Espinosa, J. F.; Frank, M.; Gilleron, M.; Kaltner, H.; Kozár, T.; Bovin, N. V.; von der Lieth, C.-W.; Vliegenthart, J. F. G.; Jiménez-Barbero, J.; Gabius, H.-J. *ChemBioChem* **2000**, *1*, 181.
- (45) Reyes-López, C. A.; Hernández-Santoyo, A.; Pedraza-Escalona, M.; Mendoza, G.; Hernández-Arana, A.; Rodríguez-Romero, A. *Biochem Biophys Res Commun* **2004**, *314*, 123.
- (46) Chávez, I., UAM, 2005.
- (47) Braunschweiler, L.; Ernst, R. R. *J Magn Reson* **1983**, *53*, 521.
- (48) Neuhaus, D. W., M. P. *The Nuclear Overhauser Effect in Structural and Conformational Analysis*; VCH: New York, 1989.
- (49) M. J. Frisch; G. W. Trucks; H. B. Schlegel; G. E. Scuseria; M. A. Robb; J. R. Cheeseman; V. G. Zakrzewski; J. A. Montgomery Jr.; R. E. Stratmann; J. C. Burant; S. Dapprich; J. M. Millam; A. D. Daniels; K. N. Kudin; M. C. Strain; O. Farkas; J. Tomasi; V. Barone; M. Cossi; R. Cammi; B. Mennucci; C.

- Pomelli; C. Adamo; S. Clifford; J. Ochterski; G. A. Petersson; P. Y. Ayala; Q. Cui; K. Morokuma; P. Salvador; J. J. Dannenberg; D. K. Malick; A. D. Rabuck; K. Raghavachari; J. B. Foresman; J. Cioslowski; J. V. Ortiz; A. G. Baboul; B. B. Stefanov; G. Liu; A. Liashenko; P. Piskorz; I. Komaromi; R. Gomperts; R. L. Martin; D. J. Fox; T. Keith; M. A. Al-Laham; C. Y. Peng; A. Nanayakkara; M. Challacombe; P. M. W. Gill; B. Johnson; W. Chen; M. W. Wong; J. L. Andres; C. Gonzalez; M. Head-Gordon; E. S. Replogle; J. A. Pople 1998.
- (50) Koradi, R.; Billeter, M.; Wüthrich, K. *J Mol Graph* **1996**, *14*, 51.

Conclusiones



CONCLUSIONES:

1. La combinación de métodos de RMN con cálculos de mecánica molecular de los compuestos **1** y **2** ha permitido demostrar que las preferencias conformacionales del *O*-disacárido (**1**) y del *C*-disacárido (**2**) son muy similares en torno al enlace glicosídico, mientras que existen diferencias en torno al enlace aglicónico. De hecho, el conformero *anti*- ψ aparece altamente poblado en el *C*-disacárido, mientras que el conformero *syn*- ψ es el mayoritario en el *O*-disacárido. Este es un comportamiento semejante al observado en otras parejas de *O/C*-disacáridos.

2. Los experimentos de NOE transferido de **1** y **2** con la galectina-1 humana indican que la conformación *syn*- ψ es la reconocida por esta lectina en ambos casos. Este hecho demuestra que el sitio de reconocimiento de la galectina-1 posee una arquitectura que permite acomodar selectivamente la conformación *syn*- ψ , independientemente del tipo de enlace entre los monosacáridos ($\beta 1 \rightarrow 4$ ó $\beta 1 \rightarrow 3$) o la naturaleza química del puente entre las unidades de monosacáridos (oxígeno o CH₂).

3. Los NOEs observados para el trisacárido Lewis a, descrito como mono-conformacional en estudios previos (conformero *syn*- $\psi(+)$ -*syn*- $\psi'(+)$), han puesto de manifiesto la existencia de una pequeña población de un segundo conformero, *syn*- $\psi(+)$ -*syn*- $\psi'(-)$, no detectado anteriormente, lo que es consistente con los estudios de mecánica molecular.

4. El estudio de RMN del análogo *C*-Le^a han mostrado la existencia de dos familias conformacionales en disolución para este glicomimético, el conformero mayoritario (*syn*- $\psi(+)$ -*syn*- $\psi'(+)$) y un conformero minoritario (*syn*- $\psi(+)$ -*anti*- ψ'), que no estaba poblado en el compuesto natural. Por tanto, la sustitución del oxígeno puente entre las unidades de galactosa y *N*-acetilglucosamina prácticamente no altera el comportamiento conformacional en torno a esta unión, pero sí provoca un aumento de flexibilidad en torno a la unión

interglicosídica contigua, entre las unidades de fucosa y *N*-acetilglucosamina. Este resultado pone de manifiesto que la importancia de los factores estéricos para regular el comportamiento conformacional de estas moléculas.

5. Mediante la aplicación del método JAMFIS, usando las constantes de acoplamiento vecinales protón-protón interglicosídicas como único dato experimental, se ha realizado el análisis conformacional de varios *C*-disacáridos previamente estudiados por otros autores. Los resultados obtenidos han demostrado que, en general, los *C*-disacáridos son más flexibles que los correspondientes *O*-disacáridos, en contra de lo descrito anteriormente por dichos autores. De hecho, se han detectado poblaciones importantes de conformeros *anti-ψ* y *non exo*-anoméricos, ausentes en *O*-disacáridos.

6. Finalmente, el estudio de las estructuras tridimensionales de los mutantes F18W y F18Nal del Ac-AMP2, mediante RMN, indica que ambos dominios presentan plegamientos bien definidos de tipo heveína, no observándose ningún cambio conformacional importante respecto a la estructura tridimensional nativa del AcAMP-2. Por tanto, estos polipéptidos son modelos muy adecuados para estudiar la importancia de la naturaleza química de los anillos aromáticos de las lectinas en el reconocimiento proteína-carbohidrato. Las estructuras de las dos proteínas mutantes libres y en presencia de (GlcNAc)₃ son similares, sugiriendo una preorganización del sitio de reconocimiento de quitoooligosacáridos. Además, los datos termodinámicos de asociación la proteína-carbohidrato revelaron un aumento de afinidad al aumentar la superficie aromática del residuo en posición 18, lo que demuestra la importancia de las interacciones tipo CH- π en los procesos de reconocimiento molecular entre proteínas y carbohidratos.

Esta Tesis ha dado lugar a las siguientes publicaciones:

- On the Importance of Carbohydrate-Aromatic Interactions for the Molecular Recognition of Oligosaccharides by Proteins: NMR Studies of the Structure and Binding Affinity of AcAMP2-Like Peptides with Non-Natural Naphthyl and Fluoroaromatic Residues. *Chem. Eur. J.* **2005**, 11, 7060 – 7074
- Hevein domains: an attractive model to study carbohydrate–protein interactions at atomic resolution. *Adv. Carbohydr. Chem. Bi.*, **2006**, 60, 303-354
- Conformational behaviour of glycomimetics: NMR and molecular modelling studies of the C-glycoside analogue of the disaccharide methyl β -D-galactopyranosyl-(1 \rightarrow 3)- β -D-glucopyranoside. *Carbohydrate Research* 342, **2007**, 342, 1910–1917

El resto de los resultados se publicarán más adelante.

